GenCore version 5.1.6

Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 21:48:11 ; Search time 7519 Seconds
(without alignments)
10742.724 Million cell updates/sec

Title: US-09-937-059-57
Perfect score: 1421
Sequence: 1 gaaccaggaaagtaaccggct.........cttctcttcttgttcttt 1421

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 2842172S653 residues

Total number of hits satisfying chosen parameters: 11766282

Post-processing: Minimum Match 0%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :
```
GenEmbl:*
 1: gb_ba:*
 2: gb_in:*
 3: gb_env:*
 4: gb_om:*
 5: gb_ov:*
 6: gb_pat:*
 7: gb_ph:*
 8: gb_pr:*
 9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1413	99.4	1441	6	AX035346	AX035346 Sequence
2		1010.6	71.1	1740	8	HSM806241	BX537581 Homo sapi
3		741	52.1	741	6	CQ727948	CQ727948 Sequence
4		691.4	48.7	1354	6	BC031111	BC031111 Mus muscu
5		626.2	44.1	5851	4	AF547266	AF547266 Sus scrof
6		453	31.9	498	8	AX035348	AX035348 Sequence
7		413.4	29.1	166937	14	AC012255	AC012255 Homo sapi
8		413.4	29.1	181312	8	AC091117	AC091117 Homo sapi
c 9		411.8	29.0	155534	14	AC009700	AC009700 Homo sapi
c 10		411.8	29.0	171444	14	AC087790	AC087790 Homo sapi
c 11		391	27.5	2684	8	AX714511	AX714511 Sequence
12		391	27.5	2684	8	AK056896	AK056896 Homo sapi
13		391	27.5	155534	14	AC009700	AK405919 Sequence
14		356.4	25.1	1636	8	AX405919	AX405919 Sequence
15		356.4	25.1	1923	8	BC029819	BC029819 Sequence
16		354.8	25.0	1029	6	AR564559	AR564559 Sequence
17		354.8	25.0	2133	6	AR564558	AR564558 Sequence
18		348.8	24.5	1521	9	BC019755	BC019755 Mus muscu
19		299.4	21.1	2439	8	AK027187	AK027187 Homo sapi
20		296	20.8	864	6	CQ727947	CQ727947 Sequence
21		245.8	17.3	1706	8	BC020841	BC020841 Homo sapl
22		245.8	17.3	2684	8	AX714511	AX714511 Sequence
23		245.8	17.3	2684	8	AK056896	AK056896 Homo sapi
c 24		216	15.2	217	8	HS183H12F	Z57552 H.sapiens C
25		199.8	14.1	218	8	HS183A12R	Z59954 H.sapiens C
26		190.4	13.4	703	10	BV621223	BV621223 S217P6048
27		190	13.4	277	10	BV188008	BV188008 sqnm15639
c 28		180.8	12.7	1338	5	BC075555	BC075555 Xenopus 1
c 29		167.8	11.8	229583	9	AL845566	AL845566 Mouse DNA
30		166	11.7	1923	8	BC029819	BC029819 Homo sapi
31		165.8	11.7	231868	14	AC118124	AC118124 Rattus no
32		159.4	11.2	166937	14	AC012255	AC012255 Homo sapi
33		159.4	11.2	171444	14	AC087790	AC087790 Homo sapi
c 34		159.4	11.2	181312	8	AC091117	AC091117 Homo sapi
c 35		154.2	10.9	231868	14	AC118124	AC118124 Rattus no
c 36		147	10.3	229583	9	AL845566	AL845566 Mouse DNA
37		123.2	8.7	110250	14	AC138220	AC138220 Mus muscu
38		121.6	8.5	55061	14	AC091597	AC091597 Mus muscu
c 39		101	7.1	110250	14	AC138220	AC138220 Mus muscu
c 40		88.2	6.2	110000	14	AC152183_2	Continuation (3 of
41		65.6	4.6	1602	6	CQ601302	CQ601302 Sequence
42		65.6	4.6	1808	6	CQ581145	CQ581145 Sequence
43		65.6	4.6	2207	2	AY061346	AY061346 Drosophil
44		65.6	4.6	3020	2	BT011438	BT011438 Drosophil
45		63.4	4.5	125020	8	AF429315	AF429315 Homo sapi

ALIGNMENTS

```
RESULT 1
AX035346
LOCUS       AX035346               1441 bp    DNA        linear   PAT 15-NOV-2000
DEFINITION  Sequence 1 from Patent WO0053748.
ACCESSION   AX035346
VERSION     AX035346.1  GI:11191064
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Bruck,C.E., Vinals,Y.D., Coche,T. and Cassart,J.P.
  TITLE     Novel compounds
  JOURNAL   Patent: WO 0053748-A 1 14-SEP-2000;
            BRUCK CLAUDINE ELVIRE MARIE (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ;
            VINALS Y DE BASSOLS CARLOTA (BE) ; COCHE THIERRY (BE) ; CASSART
            JEAN POL (BE)
FEATURES             Location/Qualifiers
     source          1..1441
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

Query Match 99.4%; Score 1413; DB 6; Length 1441;
Best Local Similarity 100.0%; Pred. No. 5.9e-237;
Matches 1413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         9 AAAGTAACGGCTACAGACAGCAGTGAGAAATAGTTTCGCTCGCGCGGCTAGAAAAACTCTGTCG     68
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 AAAGTAACGGCTACAGACAGCAGTGAGAAATAGTTTCGCTCGCGCGGCTAGAAAAACTCTGTCG     60

Qy        69 GTACCAACCCCCAGAGAGGTTGAGAGCAGCAGCCCACCACCTTCCTTAACGGAGAGAGGTGCA    128
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GTACCAACCCCCAGAGAGCGTTGAGAGCAGCCCACCACCTTCCTTAACGGAGAGAGGTGCA    120

Qy       129 GGACTCAGACTTCACCAGCCACCCACTCGGTCCGCCGCCGCTAGGATCTCCGTCCCAGCCA...    188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GGACTCAGACTTCACCAGCCACCCACTCGCTCCGTCCCGCTCGGCCGCTAGGATCTCCGTCCCAGCCA...    180
```

```
Qy    189 GCGGCTCTCCCGCGGGTCCAGGCAGCCAGCCCA 248
          ||||||| ||||||||||||||||||||||||||
Db    181 GCGCTCTTCCCGCGGGTCCAGGCAGCCAGCCCA 240

Qy    249 ACTCGGCGCCGGCGTGCAGCATGACCCTGTCCT 308
          ||||||||||||||||||||||||||||||||||
Db    241 ACTCGGCGCCGGCGTGCAGCATGACCCTGTCCC 300

Qy    309 CGGCCATGCCGCAGGCTTCAGCGGTTCACCTGCTCATCGTTATTCTAGTGTTTTTGGCTCTA 368
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CGGCCATGCCGCAGGCTTCACGGTTCACCTGCTCATCGTTATTCTAGTGTTTTTGGCTCTA 360

Qy    369 GCAGCAAGCTTCCTGCTCATCTTGCCGGGATCCGGCCACTCGGCGTGGTTGGTTG 428
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCAGCAAGCTTCCTGCTCATCTTGCCGGGATCCGGCCACTCGGCGTGGTTTTGGTTG 420

Qy    429 GTGAGAGTTCTTCAGTGTTCTGTTGTTGGCGTGTGTGCACTTCAGTGCA 488
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGAGAGTTCTTCAGTGTCTGTTCATAGGCGCAGAAATTGTGGCTGCACTTCAGTGCA 480

Qy    489 GAATGGTTCGTGGGTGTACAGTGAACACACCAACACATCCTACACAAAGCCTTCAGCCGAGCCGC 548
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAATGGTTCGTGCAGTGACACTGAACACACCAACACATCCTACACAAAGCCTTCAGCCGCGCC 540

Qy    549 GTTACAGCCCGTGTCGGTGTTGCGTGGTGGGCCTGGAGGGGCCATTAATATTACACTCACAGGG 608
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTTACAGCCCGTGTCGGTCTGTCGTGGGCCTGGAGGGGCATTAATATTACACTCACAGGG 600

Qy    609 ACCCCAGTGCATCAGCTGAACGAGACGACCATTGACTACTACAACGAGAGCAGTTCACCTGGCTCTG 668
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACCCCAGTGCATCAGCTGAACGAGACCATTGACTACTACAACGAGAGCAGTTCACCTGGCTCTG 660

Qy    669 AAAGAGAATTACCGCCCGCGAGTACGGCGAACGCACTGGAGAAGGGCTGCGCCGGACCCAGTG 728
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAAGAGAATTACCGCCCGCGAGTACGGCGAACGCACTGGAGAAGGGCTGCGCCGGACCCAGTG 720

Qy    729 CTCTACCTGGCGGAGGAAGTTCACCGAGTAGCCCTTGCCCGTGTACCACCAGTACCAC 788
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCTACCTGGCGGAGGAAGTTCACACGGAGTAGCCCTTGGGCCGTGTACCACCAGTACCAC 780

Qy    789 CTGGCGGGGACACTAGCCTCGGCGCCACGCTATGGGGTTGGGTTCTTGCTTCTGGGCTCTCC 848
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CTGGCGGGGACACTAGCCTCGGCGCCACTAGCCGTTCTGCTTCTGGCTCTCTCTC 840

Qy    849 AACGTGCTGCTTCTCTCACGCGCCGGGCCCCGCTCTACGGGAGGGGCCTGGCACTGCTACCACCGGGA 908
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACGTGCTGCTTCTCTCACGCGCGCCCCGCTCTACGGGAGGGGCCTGGCACTGCTACCACCGGA 900

Qy    909 GCCTTCGCGCCTTCGGGGTTCGCCTTGGCCTTCCATCTTCTAGGGTGCGCGCTCTGCCCCG 968
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GCCTTCGCGCCTCTTCGGGGTTCGCCTTGGCCTTCCATCTTCTAGGGTGCGCGCTCTGCCCCG 960

Qy    969 CTCCGCCCTAGGCTCCTCCGGGCGCTCACCACTCAGTACGGCGGGCCCTTCGGGTCAGGCTG 1028
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CTCCGCCCTAGGCTCACCACTCAGTACGGCGGGCCCTTCGGGTCAGGCTG 1020

Qy   1029 GCAACCGGGCGTCCTTGTGCCCTCTTCCTCGGAGGGGGGGCGCGGGGGCCGTGGTGTGAGTCTTCAGTCT 1088
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCAACCGGGCGTCCTTGTGCCCTCTTCCTCGGAGGGGGGCGGGGGCCGTGGTGTGAGTCTTCAGTCT 1080

Qy   1089 CCCAGCGGCTCTTCGGCACCCTTCTCGGACCAAAAGCGCGCAAGGAACTGCGAGCCAGGAGAGG 1148
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CCCAGCGGCTCTTCGGCACCCTTCTGGACCAAAAGCGCGCAAGGAACTGCGAGCCAGGAGAGGG 1140

Qy   1149 GGCTCACCTCTTATCCTCGGGGACCACTGCACAAGCAGAGCGCTCTCCCAGGACTTAAAA 1208
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GGCTCACCTCTTATCCTCGGGGACCACTGCACAAGCAGAGCGCTCTCCCAGGACTTAAAA 1200

Qy   1209 TGTATCACCACTAACCTGTGAGGGGGACCCAAATCTTGGGACTCAGCAT 1268
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TGTATCACCACTAACCTGTGAGGGGGACCCAAATCTTGGGACTCAGCAT 1260
```

```
Qy   1269 CGGAGGGCCGGGGAAGCAGTGCCCCCAGCCCCAGCCCCAGGAGGAGAGGGGCACTG 1328
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCAGGGGCGGGGAAGCAGTGCCCCGCGCCAGGCGCCTGGGCCCAGGAGAGTTCCAGGAAGGGGCACTG 1320

Qy   1329 AGCGCTGTGGCGCGGAGCCTCGGACAATCCGACAGCACACGCAGCAGGAAAAGTCTTCCTGGGGGGA 1388
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGCGGTGTGGCGCGGAGCCTCGGACAATCCGACAGCAGCACGCAGCAGGAAAAGTCTTCCTGGGGGGA 1380

Qy   1389 TCTGTAAATAAACCTTTTTTTTTCTTTTGTTTTTTT 1421
          ||||||||||||||||||||||||||||||||||||
Db   1381 TCTGTAAATAAACCTTTTTTTTTCTTTTGTTTTTTT 1413
```

RESULT 2

HSM806241

```
LOCUS       HSM806241              1740 bp    mRNA    linear   PRI 17-JUN-2003
DEFINITION  Homo sapiens mRNA; cDNA DKFZp686C04213 (from clone DKFZp686C04213).
ACCESSION   BX537581
VERSION     BX537581.1  GI:31873659
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1740)
  AUTHORS   Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            Clone from GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            This clone (DKFZp686C04213) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY, Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1..1740
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="15q15"
                     /clone="DKFZp686C04213"
                     /tissue_type="human rectum tumor"
                     /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
                     DH10B; sites SfiIA + SfiIB"
                     /dev_stage="adult"
     polyA_signal    1694..1699
     polyA_site      1720
ORIGIN
```

```
Query Match             71.1%;  Score 1010.6;  DB 8;  Length 1740;
Best Local Similarity   98.6%;  Pred. No. 1.1e-166;
Matches 1019;  Conservative 0;  Mismatches 14;  Indels 0;  Gaps 0;
```

```
Qy    389 CTTGCCGGGATCCGGCCACTCGGCGTGGTTTTGGTTGGTGAGAGTTCTTCAGTCT 448
          |||   |||        ||||||||||||||||||||||||||||||||||||||
Db    688 CTTGACCCGGGCGTGCCTATTCCTGCAGCGCGTGGTTTTGGTTGGTGAGAGTTCTTCAGTCT 747

Qy    449 GTTCATAGGCGCAGAAATTGTGGCTGCACTTCAGTGCACGTGTGCAGAATGGTTCGTGGGTGTACAGT 508
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    748 GTTCATAGGCGCAGAAATTGTGGCTGCACTTCAGTGCAGAATGGTTCGTGGGTACAGT 807

Qy    509 GAACACACCAACACATCCTACACAAAGCCTTCAGCGCGCGCCGGGTTTACAGCCCGTGTCGGTCT 568
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    808 GAACACACCAACACATCCTACACAAAGCCTTCAGCGCGCGCCGGGTTACAGCCCGTGTCGTGT 867

Qy    569 GCTCGTGGGCCTGGAGGGGCATTAATATTACACTCACAGGGACCCCAGTGCATCAGCTGAA 628
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    569 GCTCGTGGGCCTGGAGGGGCATTAATATTACACTCACAGGGACCCCAGTGCATCAGCTGAA 628
```

```
RESULT 3
CQ727948
LOCUS       CQ727948               741 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 13882 from Patent WO02068579.
ACCESSION   CQ727948
VERSION     CQ727948.1  GI:42295393
KEYWORDS
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 13882 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1..741
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match        52.1%;  Score 741;  DB 6;  Length 741;
Best Local Similarity   100.0%;  Pred. No. 1.6e-119;
Matches  741;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy   485 TGCAGAAATGGTTCGTGGGGTACAGTGAACACCAACACATCCTACAAAAGCCTTCAGCGCAGC 544
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TGCAGAAATGGTTCGTGGGGTACAGTGAACACCAACACATCCTACAAAAGCCTTCAGCGCAGC 60

Qy   545 GCGCGTTACAGCCCCGTGTCGTCTCGTGGGGGTCGGGAGGGCGCATTAAATATTACACTCAC 604
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GCGCGTTACAGCCCCGTGTCGTCTCGTGGGGGTCGGGAGGGCGCATTAAATATTACACTCAC 120

Qy   605 AGGGACCCCAGTGCATCAGCTGAACGGAGACCAATTGACTACTACAACGGACCAGTTCACCTGGCG 664
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGGGACCCCAGTGCATCAGCTGAACGGAGACCAATTGACTACTACAACGGACCAGTTCACCTGGCG 180

Qy   665 TCTGAAAGAGAATTACGGCCGGGGAGAAGTTCACACCGGAGGAAGGGCTGGACTGACCAC 724
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TCTGAAAGAGAATTACGGCCGGGGAGAAGTTCACACCGGAGGAAGGGCTGGACTGACCAC 240

Qy   725 AGTGCTCTACCTGGCGGAGAAGTTCACACCCCGAGTAGCCCTTGCGGCGCCTGTACCCAGTA 784
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AGTGCTCTACCTGGCGGAGAAGTTCACACCCCGAGTAGCCCTTGCGGCGCCTGTACCCAGTA 300

Qy   785 CCACCTGGCGGAGCACTACGGCCTCGGCCACCCTATGGGTTGCGTTCTGCTTCTGGCTCCT 844
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCACCTGGCGGAGCACTACGGCCTCGGCCACCCTATGGGTTGCGTTCTGCTTCTGGCTCCT 360

Qy   845 CTCCAACGTGCTGCTCTCCAGCCGCCGGCCGCCCGGCTGTACGGGAGCCTGGCACTGACCAC 904
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTCCAACGTGCTGCTCTCCAGCCGCCGGCCGCCCGGCTGTACGGGAGCCTGGCACTGACCAC 420

Qy   905 CGGGAGCCTTGCGCGCTCTTCGGGGTCTTCGCGGTGCGCGGCCACTGCAGCCGGAGGAGAG 964
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CGGGAGCCTTGCGCGCTCTTCGGGGTCTTCGCGGTGCGCGGCCACTGCAGCCGGAGGAGAG 480

Qy   965 CCCGCTCCGGCTAGGGTCCTCCGGGGCTCACCCACTCAGTACCGGCGCGCCCCTTCTGGGTCAC 1024
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CCCGCTCCGGCTAGGGTCCTCCGGGGCTCACCCACTCAGTACCGGCGCGCCCCTTCTGGGTCAC 540

Qy  1025 GCTGGGCAACCGGCGTCCTGTGTGCCTCTTCCTCGGGAGGGGGCCGGTGGTGAGGTCTCCAGTATGT 1084
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCTGGGCAACCGGCGTCCTGTGTGCCTCTTCCTCGGGAGGGGGCCGGTGGTGAGGTCTCCAGTATGT 600

Qy  1085 TCGGGCCCAGCGGCTCTTCGCACCCTTCTGGACCAAAAGCGGCCAAGGACTGCAGCCGGAGGAGAG 1144
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TCGGGCCCAGCGGCTCTTCGCACCCTTCTGGACCAAAAGCGGCCAAGGACTGCAGCCGGAGGAGAG 660

Qy  1145 AGGGGGGCTCACCTCTTATCCTCGGCCGACCCACTGCACGCAAGCACCGGCGCTCTCCCAGACTT 1204
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AGGGGGGCTCACCTCTTATCCTCGGCCGACCCACTGCACGCAAGCACCGGCGCTCTCCCAGACTT 720

Qy  1205 AAAATGTATCACCACTAACCT 1225
             |||||||||||||||||||||
Db   721 AAAATGTATCACCACTAACCT 741
```

```
RESULT 4
BC031111
LOCUS       BC031111              1354 bp    mRNA    linear   ROD 06-JUL-2005
```

```
Qy   629 CGAGACCAATTGACTACTACAACGGACCAGTTCACCTGGCGTCTGAAAGAGAATTACGGCCGGGGA 688
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   868 GCTCGTGTGGGCCCTGGAGGGGCCATTAATATTACACTCACAGGGACCCCAGTGCATCAGCTGAA 927

Qy   689 GTACGGGAAGCCACTGGAGAAGGGGGCTGCCGGAGGACCCAGTGCTCTACCTGGCGGAGGAAGTT 748
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   928 CGAGACCAATTGACTACTACAACGGACCAGTTCACCTGGCGTCTGAAAGAGAATTACGGCCGCGA 987

Qy   749 CACACCGGAGTAGCCTGCGTGTACCAGTACCACCTGGCGGAGCACTACGGCCTCGGCCACCCTC 808
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   988 GTACGGGAAGGCACTGGAGAAGGGGGCTGCGGAGGACCCAGTGCTCTACCTGGCGGAGGAAGTT 1047

Qy   809 GGCCACGCTATGGGTGGGCTTCTGCTTCTGGCTCCTCTCCAACGTGCTGCTCTCCAGCCGCC 868
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1048 CACACCGAGTAGCCCTGTACCACGTACCACCAGTACCACCTGGCGGAGCACTACGGCCTC 1107

Qy   869 GGCCCCGGCTTACGGGAGGCCTGGCACTGACCACCGGGAGCCTTCGCGCTCTTCGGGGGT 928
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1108 GGCCACGCTATGGGTGGGCTTCTGCTTCTGGCTCCTCTCCAACGTGCTGCTCTCCCAGCC 1167

Qy   929 CTTCGCGGTGCGCGGCCACTGCAGCCGGAGGAGAGCCCGCTCCGGCTAGGGTCCTCCGGC 988
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1168 GGCCCGGCTTACGGGAGGCCTGGCACTGACCACCGGGAGCCTTCGCGCTCTTCGGGGGT 1227

Qy   989 GCTCACCCACTCAGTACCGGCGCGGGGTCTTCGGCGTCGGGGTCTTCGGCCGTTCCTGTCCT 1048
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1228 CTTCGCGGTGCGCGGCCACTGCAGCCGGAGGAGAGCCCGCTCCGGCTAGGGTCCTCCGGC 1287

Qy  1049 CTTCCTCGGAGGGGGCGTGGGTGAGTCTCCAGTAGTGTTCGGCCCAGCGGCTCTTATCCTCCGACCCT 1108
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1348 GCTCACCCACTCAGTACCGGCGCGCCCCGTTCCAGCGGCTTCCAGTGCAACCGGGCTCCGTGCCT 1347

Qy  1109 TCTGACCAAAGCGGCCAAGGACTGCAGCCGCGAGCTCTGGGGACATCGCAGGCAGGCAGTGC 1168
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1408 CTTCCTCGGAGGGGGCGTGGGTGAGTCTCCAGTAGTGTTCGGCCCAGCGGCTCTTATCCTCGG 1467

Qy  1169 CGACCCACTGCACAAGCAGCGGCCGCCGTCTCCAGACTTAAAAATGTATCACCACTAACCTGTGTG 1228
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1468 CGACCCACTGCACAAGCAGCGGCCGCCGTCTCCAGACTTAAAATGTATCACCACTAACCTGTG 1527

Qy  1229 AGGGGGGACCCAATCTGGACTCCTTCCCCGGCCTTGGGGACACCGGGGGAAGCAGTGC 1288
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1528 AGGGGGGACCCAATCTGGACTCCTTCCCCGGCCTTGGGGACACGGGGGGAAGCAGTGC 1587

Qy  1289 CCGCCAGGCCTGGGCGCCAGGAGAGCGCTCCAGGAGAGCTTCCAGGAAGGGCCACTGAGGCTTCTGGGTGCTGCTG 1348
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1588 CCGCCAGGCCTGGGCGCCAGGAGAGCGCTCCAGGAGAGCTTCCAGGAAGGGCCACTGAGGTGCTGCTG 1647

Qy  1349 TCGGACAATCCGCAGCCACCAGGGAAAGTCTTCTCCGAGAAGGGCCACTCAGTACCGGCGCGCTGTGT 1408
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1588 CCGGCCAGGCCTGGGCGCCAGGAGAGCTTCCAGGGTCTTGCTCTGGGTGCTGCTG 1647

Qy  1409 TCTTTTGTTTTTTT 1421
             ||||||||||||||
Db  1708 TCTTTTGTTTTTTT 1720
```

```
DEFINITION  Mus musculus RIKEN cDNA 9030623N16 gene, mRNA (cDNA clone MGC:35715
ACCESSION   BC031111
VERSION     BC031111.1  GI:21411407
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1354)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED    12477932
REFERENCE   2  (bases 1 to 1354)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site: http://www-shgc.stanford.edu
            Contact: (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
            R. M.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 58 Row: g Column: 17
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 13385245.
FEATURES             Location/Qualifiers
     source          1..1354
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:35715 IMAGE:4989815"
                     /tissue_type="Colon, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1..1354
                     /gene="9030623N16Rik"
                     /db_xref="GeneID:66811"
                     /db_xref="MGI:1914061"
```

```
CDS             129..1091
                /gene="9030623N16Rik"
                /codon_start=1
                /product="Numb-interacting protein 2"
                /protein_id="AAH31111.1"
                /db_xref="GI:21411408"
                /db_xref="GeneID:66811"
                /db_xref="MGI:1914061"
                /translation="MTAWDGVLPFYPQPRHAASFSVPLLIVILVFLSLAASFLFILPG
                IRGHSRWFWLVRVLLSLFIGARSIVAVHFSGDWFGVGRVWTNTSYKAFSPSRVQVHGLH
                VGLAGVNITLRGTPRQQLNEFIDYNRFWTFWRLNEBDYTKEYVHALEKGLPDPVLYLARK
                FTPSSPCGLYHQYHLAGHYAAATLWVAFCFWIIANALLSMPAPLYGGLALLTTGAFTL
                FGVFAPASISSVPLCHFRLGSAVLIPYYGASFWLTLATGILSLLLGGAVVILHYTRPS
                ALRSFLDLSVRDCSNQAKGNSPLITLNNPQHEQLKSPDLNITTLL"
```

```
ORIGIN
        Query Match     48.7%;  Score 691.4; DB 9;  Length 1354;
        Best Local Similarity   75.8%;  Pred. No. 6.9e-111;
        Matches  913;  Conservative   0;  Mismatches  276;  Indels  16;  Gaps   4;

Qy        230 CTGCCCGCGCCGCGTGCAGCACTCGGCGCGGCCGGCGTGCGGCGACGCCGTGTGGAAACGGCGTACT  289
              |||||||||||||  ||| |||||||| |||||||||||  || |||| |||||  | |||| ||||
Db         92 CTCGCCTGCCTGCGAACAGTTCTCGAACTTCTTGCAGGATGAGACTGCGTGGGACGCGGGTGCT  151

Qy        290 GCCTTTTACCCCCGGACACCCCCGGACGCCCGCAGGGCTTCAGCGGTTCCACTGCTCATCGTTAT  349
              |||||||||||||  ||||||||||||  ||||||| | |||||||||||||| | ||||| |||
Db        152 ACCCTTTTACCCCCAGGCCATGCCCCCAGCCGCGGTTCCAGCGTACCGCGCTGCTCATTGTTAT  211

Qy        350 TCTTAGTGTTTTTTTGGCTCTAGCAGCAGCAAGCTTCCTCCTCCTCCACTTGCCGGGGGATTGT  409
              |||||||||||||||||  | |||| || |||| |||||  || |||||||| || |||||||
Db        212 CCTGGTGTTTCTTGTCTGTAGCTTGCTGTCTAGCTTGCATCTTGCCTGGAATCCGTGGCCA  271

Qy        410 CTCGCGGCCTGGGTTTTGGTTGGTGGAGAGTTCTTCTTCAGTCTGTTCATCAGGGGCAGAAATTGT  469
              |||||||||||||||||||  ||| |||||| || |||| |||| |||| | ||||  ||| |||
Db        272 CTCGGCCTGGTTCTCTGGTTGGTAAGAGTCCTTCTTAGCTCTTGCATAGGTGCCGAAATTGT  331

Qy        470 GGCGTGCACTTCAGTGCAGAATGGTTCGTGGTGTACAGTGAAACACCAACACATCCTACAA  529
              |||||||  || ||| || |||||||| |||||| ||||| ||| ||||||| ||||||||
Db        332 GGCTGTGCACTTCAGTGGACTGGTTCGTGCTGGAGAGTGTGGACCACAACATCCTACAA  391

Qy        530 AGCCTTCAGCGGACAGCGCGCGTTACAGCCCGTGTCTGTCGGTCTGCTCGTCGGTGGGGCAT  589
              |||||||||||  |||||||||||||| |||  |||||||||| ||||||| | |||||||
Db        392 AGCCTTCAGTCATCCACGTGTTCAAGTCCATGTGCGGTCTCGCACTGCTGGGCGCGGGGCGT  451

Qy        590 TAATATTACACTCACAGGGAGACCCCAGTGCATCAGCTGAACAGGAACACCAGTTGACTACAACGA  649
              ||| || || | |||||||| ||||||||| |||| |  || || | || |||||||| ||
Db        452 TAACATTACACTCCCAGGACACCGAGCAGCTGAACGGAGACCATTGACTACAACATGA  511

Qy        650 GCAGTTCACCTGGCGGTTCTGAAAGAGAATTACGCCCGGGACTACGGGACCACTGGAGAA  709
              |||||||||  |||||||||||| | |||  ||| |||||  || ||| ||||||||||
Db        512 GCAGTTCACTTGGCGGTCTGAACGAGACTACACCAAGGAGTATGTCCATGCCTTGGAGAA  571

Qy        710 GGGGCTGCCGGACCCCAGTGCTCTACCTGGCGGAGAAGTTCACACCGAGTAGCCCTTGCGG  769
              ||||  || |||  ||||||  | ||||| || ||||| ||||||||||||  |||||||
Db        572 GGGACTGCCGGACCCCAGTGCTTATACCTGGCAGAGAGCAGTCCTTGCGG  631

Qy        770 CCTGTACCACCAGTACCACCTGGGCGGGACACTACCCCCAGCTCAGGAGGCCT  829
              |  ||| || ||||| |||||| |||  ||||||||  |||| |||||||||
Db        632 GCTGTACCACCAATATCACCTGGCCGGGTCACTATGCGCAAACACTGTGGGGTGGCATT  691

Qy        830 CTGCTTCTGGGCTCCTCTCCAACGTGCTGCTCTCCAACGCCGCGGCCCCGCCCTTACGGAGGCCT  889
              |||  |||||||| ||  |||||||| || |||||| || ||||  |||||  |||||||||
Db        692 CTGCCTTCTGGATCACCGCCAATGCCGCTCTCCAGATGCCCGCCCCGCTTTACGGAGGCCT  751

Qy        890 GGCACTGCTGACCACCGGAGCCTTCGCGCTCTTCGGGGTCTTCGCCCTTGGCCTCCATCTC  949
              |||||||||| || ||||| | |||| |  | |||| |||||||||| |||| |||||
Db        752 GGCTTTGCTCACCACCGGCCGGCTTCACGGCCTTCGTGCTTTCGCGCTGGGCATTCGATTTC  811

Qy        950 TAGCCGTGGCCCCCGCTCTCGCCCGGCGTCTCCGGGGCTTCACCACTCAGTACCGGCG  1009
              |||||||||||||  | ||||  |||||||||||| ||||||||||| ||||||||||
Db        812 CAGCGGTTCCGGCTGTCCACTTCCGCCCTTGGGCTCCACCACCCTTACTACGGGCGC  871

Qy       1010 CGCCTTCTGGGGTCACGGCTGGCGAACGGCAGTCCGTTCCTCTTCGGGAGGGGCGGTGGT  1069
```

RESULT 5

```
LOCUS       AF547266                5851 bp    mRNA    linear   MAM 03-APR-2003
DEFINITION  Sus scrofa dual oxidase 1 (Duox1) mRNA, complete cds.
ACCESSION   AF547266
VERSION     AF547266.1  GI:23664372
KEYWORDS    .
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 5851)
  AUTHORS   Morand,S., Chaaraoui,M., Kaniewski,J., Deme,D., Ohayon,R.,
            Noel-Hudson,M.S., Virion,A. and Dupuy,C.
  TITLE     Effect of iodide on nicotinamide adenine dinucleotide phosphate
            oxidase activity and Duox2 protein expression in isolated porcine
            thyroid follicles
  JOURNAL   Endocrinology 144 (4), 1241-1248 (2003)
  PUBMED    12639906
REFERENCE   2  (bases 1 to 5851)
  AUTHORS   Kaniewski,J., Morand,S., Noel-Hudson,M.-S., Ohayon,R., Virion,A.
            and Dupuy,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-SEP-2002) Unite 486, INSERM University of Paris-Sud,
            5 rue JB Clement, Chatenay-Malabry 92296, France
FEATURES             Location/Qualifiers
     source          1..5851
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
     gene            1..5851
                     /gene="Duox1"
     CDS             948..5609
                     /gene="Duox1"
                     /EC_number="1.6.3"
                     /function="putatively generates H2O2 by transferring
                     electrons from NADH/NADPH to molecular oxygen"
                     /function="oxidoreductase acting on NAD(P)H with oxygen as
                     acceptor"
                     /note="NADPH:O2 oxidoreductase"
                     /codon_start=1
                     /product="dual oxidase 1"
                     /protein_id="AAN39338.1"
                     /db_xref="GI:23664373"
                     /translation="MGPRLALAWTLLVGPWMPMGAARNSISWEVQRFDGWYNNLMBHKW
                     GSKGSRLQRLVPASYADGVYQPLGEPHILPNPRDLSNTAMRGPAGQASLRNRTVLGVFP
                     GYHVLSDLVSIERPGCPAEFLNIHIPGDPVFDPHKSGDVFLVPQRSRWNDPNTGQSPS
                     NPRDLTNEVTGWLDGSAIYGSSHSWSDELRSFSGGQLASGPDPAFPRQQDPLFWMTP
                     PDPATGQRGPQGLYAFGAEQGNRREPFLQALGLLMFRYHNICAQKLAREHPLWGDBELF
                     QHARKRVIATYQSITTMYEMLPSFLRKMPQBYTGYRPFLDPISPEFLAASEQFSTMV
                     PGVYMRNASCHFGVIRNSSVSRALRVCNSTWSREHPNLQRAEDVDALLIGMASQI
                     ARREDHMVVEDVQDFWPGLPLKPSRTDHLASCLQRGRDLGLPSTVKARRLGLPVVTRW
                     QDINPALSRSDGIVLBATAALVNQDLSRLRLELLPGLLLRSYGDPGLPFSTIVLDQPVRL
                     RDGDRYWFENTKNGLFSEKEIAEIRNTSLRDVLVAVTNMTPGALQPNVFFWHAGDPCP
                     QPRQLSTKDLPACAPLIMRDYPKGSGFPGVTIGTLCCFPLVSLLSAWIVAQLRRNP
                     KRLQVNRQSIMCEKLVGGMKALBWOGRRKEPCRPVLVHLQSGQTIHVMDGRLSVRTIQ
                     LRPPQOVNLILSSNHGRRTLLLKIPKEYDLVLMFDLEEROVMENLQSALKBSGLSF
                     QEWLREQELMRAAVTREQRSHLLETFPRHLFSQVLDIDOADAGLPLDSSQKVREAL
                     RLMFRMYDPDGNGLISKDBFIRMLRSFIEISNNCLSAQLTEVVESMRREAGFQQDKQB
                     LTWEDFHFMLRDHDSELRFTQLCVRGVGVEVPEVIKDLCRRASYISQBKLCPSRVSAHC
                     PRSNVDVEVLITPWKLQCPTDTDPPOIRRRFRPGKKVTSFQPLLFTEAHRBKFQRSRRH
                     QTVQOFKRFVENYRRHIGCLAVFTIAGGLFLRBAYYYAPAAHHMGITDDTRVGIILS
                     RGTAASISPMFSYILLTMCRNLITFLRETFLNRVPFDAAVDFHRLIASTAIILTVLH
                     SAGHVVNVLFSISPLSVLSCLFPGLFHDNGSEFPQKYTWWFQTVPGLIGVMLLILIL
                     AIMYVPASHHFRRCSFRGFWLIHHLYILLYMLLIIHGSFALIQLPRFHFPFLVPALIY
                     VGDKLVSLSRKVEISVVRABLLPSGVTHLQOFQRPGFEYKSGQWNRIACLALGTTEY
                     HPFTLTSAPHEDTLSLIHIRAAGFMVITMLRLVGRGDGSPTLDDNCAKYPKLKYLDGPFGHGQB
                     WHKFEVSVLVGGGIGVTPFPASILKDLVFKSSVSSCQVFCKKIYFIWVTRTQRQFRWLAD
                     IIREVEENDHRDLVSHIYIITQLABKFDLRTTMLYICERHFQKVLNRSLFTGLRSITH
                     FGRPFEPFFNSLQEVHPQVRKIGVFSCGPPGMTKNVEKACQLINRQDRTHFSHYEN
                     F"
```

ORIGIN
```
Query Match           44.1%;  Score 626.2;  DB 4;  Length 5851;
Best Local Similarity   81.8%;  Pred. No. 1.4e-99;
Matches  735;  Conservative  0;  Mismatches  163;  Indels  1;  Gaps  1;
```

```
Qy     262 GCAGCAATGACCCTGTGGAACGGCGTACTGCCTTTTTACCCCCAGCCCCGGCCAATGCCGCAG 321
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      10 GCAGCAATGACTCTGTGGAACGGCGGCGGGTGCCCTTCTTTACCCCTCAGCCCTCGGTATGCCCCA 69

Qy     322 GCTTCAGCGTTCCACTGCTCATCGTTATTCTAGTGTTGTTTTTGGCTCGTACGAGTTCCTTTC 381
           |||||||||||||||| || ||||||| || ||||||||||||| || ||||||||||||| |
Db      70 GCTTCAGCGTCCACTACTCATCGTTATTCTGGTGTTCTTTGGCGCCTTAGCCACCAGCTTTC 129

Qy     382 TGCTCATCTTGGCGGGGGATCCGTGGCCACTCCGCGTCGGTTTGGGTTGGTAGAGTTCTTC 441
           ||||||| ||||||| |||| ||  ||||||| ||| |||||| |||| || ||||| |||
Db     130 TACTTCATCTTGCCCGGGATTCGTGGCCACTCGCGCGTTGGTTCGGGTTGGTGAGAGTTCTCC 189

Qy     442 TCAGTCTGTTCATAGGCGGCCAGAAATGTGTGCGTGCACTTCAGTGCAGAAATGGTTCGGTGG 501
           |||  ||||||||||||| |||||| |||||||| |||||||||| ||||||| ||||||
Db     190 TCAGCCTGTTCATTAGGCGGCCAGAAATTGGTGGCGGTGCACTTCAGCGCCGGCGAGAATGGACCCTG 249

Qy     502 GTACACGTGAACACCACACATCCTACACAAAGCCTTCAGCGCGCACGCGCGGGTTACACGCCGTG 561
           ||||||||  |||| ||| || |||||||| || |||||||||| || |||||||||| ||||
Db     250 GCGAAGTTAGCAGCAATACATCTTACAAGGCCTTCAGTGTGGCGACGTGTCCAAGCCACC 309

Qy     562 TCGGTCTGCTCCGTGGGGCCTGGAGGCCATTAATTACACTCACCAGGGAGACCCCCAGTTGCATC 621
           ||||||||| |||||||||||| ||| |||||| || || || || || |||||||||||||
Db     310 TCGGTCTGCACGTGGGCCCTGGAGGGTGTGAATATCACACTCACAGGGAATCCGGTACAGC 369

Qy     622 AGCTGAACGAGACCATTGACTACAACGAGCAGTTCACCTGGCCTCTGAAAAGAGAATTACG 681
           ||||||| ||| || ||  ||||||||||||| ||||||||| || ||||| |||||||||
Db     370 AGCTGAATGAGACCATCGACTACAACGAGCAGTTCATCTGGCCTGAAAAACTATG 429

Qy     682 CCGCCGAGTACGGCGAACGCACTGGAGAAGGGGGCTGCCGGACCCAGTGCTTACCTGCCG 741
           || ||||||| |||||| ||||| ||| |||| |||| || |||||||| || |||||||
Db     430 CCGCCGGCCGTATCGGAAGGCTCTGGAGAAGGGGCCGACCCAGTGCTCTACTTGGCGG 489

Qy     742 AGAAGTTCACACCGAGTTAGCCCTTGCGGCGTCGTTACCACCAGTACCACCTGGCGGGGGACACT 801
           |||||||  |||  ||||||||||||||| || ||||||| || |||||||||||||||||
Db     490 AGAAGTT-ACTCCCAGAAGCCCTCGCGGGGTTTACCCCCAAATATCGCCTAGCCGGAGACACT 548
```

RESULT 5

```
Db     872 CTCCTTTTTGGCTCACGCTGGCCACCGGCAATCCTGAGCCTCCTCCTCTCGGAGGGGGTGGT
           ||||   |||||   |  |  |||||||||||||||| ||||||||||||||| |||||||
                                                                    931

Qy    1070 GAGTCTCCAGTATGTTCCGCCCCAGCCGCCTCGTTCCCCACCCCTTCTCGGACAAAGCCCAAGGA 1129
           ||||| || || || || ||  || |||||||| |||||  || ||| |||||| ||  |||||
Db     932 GAATCTCCACTATACTCGCGCCCCACCGCCCCTTCGCTCCTTTTCTCGGATCTTAAGTGTCAAAGA 991

Qy    1130 CT--GCAGCCAGAGAGAGAGGGGGGTCACCTCTTCTTATCCTCGGGCCGACCCACTGCACAAGCA 1186
           ||   || |  || || || ||||| |||| ||||||||| ||| ||| |||  |||||||||
Db     992 CTGTAGCAACCAGGCCTAAAGGAAACTCACCTCACCTCTTCACCCCTCAACACGAACA 1051

Qy    1187 GGCCGGCTCTCCCACGACTTAAAATGTATCACCACCACTAACCTGTGAGGGGACCCAATCTCGGA 1246
           ||||  || ||||| |  |  ||| | ||||||||||||| |||  ||||| |||  ||| |||
Db    1052 GTTGAAGAGTCCAGACTTAAA---TATTACCACTCCTGAAAATAACTGACTCCCGGA 1108

Qy    1247 CTCCTTCCCCGCCTTGGACGATCCGCAGGCGCGGGAAGCAGTGCCGCGCCCTG-GGCCA 1305
           ||  |||||| ||||| |||| || |||| || ||| ||||| |||||||  |  ||||
Db    1109 TTTCTACCCCTCTTTTTGGCGATCCGCAGGCGCGGGACCGGTGCCAGA 1168

Qy    1306 GGAGAGCTCCAGGAAGGGACACTGACCGTGCTCTGGGCGCGGGAGGCCTCGGACATCCGGCAGGCA 1365
           |||||  || |||||||| |||||  ||| | || ||||| ||| |||| ||| ||||||||||
Db    1169 GCATAGCTCGAGTCTGGCGCTTGTCCCCTGCGCGGACGACAAGGGCGCACCCCAGTTG 1228

Qy    1366 CC-----AGGGAAAGTCTCCTGTGGGGCGGATCTGTAAATAAACCTTTTTTTTCTTTTGT 1416
           ||      |||  ||  || |||| |  |||| ||||| |  ||| |||||| ||||||
Db    1229 CCTTTGCTCTAGAGAAATGATTTCTCAGGAGAACTGTAAATAAACTTTTTGTTTTCTTTTT 1288

Qy    1417 TTTTT 1421
           |||||
Db    1289 TTTTT 1293
```

```
Qy    802 ACGGCCTTCGGCGGCCACGGCTATGGGGTGGGCGTTCTTCTGGCTCCTCCTCCAAGGTGCTGCTCT 861
          ||||| ||| || ||  ||   ||  | ||| | ||||| |||| |||| |||| || ||||||||
Db    549 ACGGCCTCCGCGCACATTGTGGGACGGGCGTTCTGCTCTTCTCCAACCTGCTGCTTT 608

Qy    862 CCACGCGCCGCGGGCCCGGCTCTGACGGAGGGCGCGGTTCCTTGCGGCTCT 921
          ||||   |||||||  ||||  | ||  || ||| |||| |||  |||||||
Db    609 CCATGCCGGGTCCGGCAATTATGGAGGCGCTCCTCATCACACCGGGGCCCTCT 668

Qy    922 TCGGGGTCTTGGCCTTGGCCTCCATCTCTACGGTGCCGCGTCCGCCTTAGGCT 981
          | |||  |||||| ||||| || ||  ||  |||| |  ||| |||| ||||||
Db    669 TCTCTGTTCTTCGGCCTTCGCTCTCTTATCTCCCAGCGGTCCGCCTTCGGCCTTCT 728

Qy    982 CCTCGCCGCTCACCCACTCACACGTACGGTACGGCGGCGCGCCGCCGCCGTCC 1041
          |||| ||| ||||| ||||| |||  | || || |||  || ||| ||||| ||
Db    729 CCTCCGAACTCACCACTCACTACGGTACGGCGTGCTGCTGCCCAGCGGCCACGTCC 788

Qy   1042 TGTTGCCTCTTCCTCCGGAGGGGGGCGTGGTGAGTCTCCAGTATATGTTCGGCCCAGGGGCTTC 1101
          || || ||||| ||| ||| |||  | | ||| |||| ||   | | ||  |||| ||||||| |
Db    789 TGTTGCCCTCCTCCTCGGACGGCGCGGTCCCACTCACCTACGCTGGACACGCGCTGCTGGCCCAGGGGCTCTTC 848

Qy   1102 GCACCCTTCTGGACCAAAGCGCCAAGGACTGCAGCCAGGAGGAGGGGGGCTCACCTCTT 1160
          |||| |||||| || || || || ||| | |||| |||| || || || |||  |||| |
Db    849 GCACGTTCTTAGAGAGGAAGTAATCAAGGACTGCGGATAGTCGGATAAAAGGGTCTCCATTT 907
```

RESULT 6
```
LOCUS       AX035348                 498 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 3 from Patent WO0053748.
ACCESSION   AX035348
VERSION     AX035348.1  GI:11191065
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Bruck,C.E., Vinals,Y.D., Coche,T. and Cassart,J.P.
  TITLE     Novel compounds
  JOURNAL   Patent: WO 0053748-A 3 14-SEP-2000;
            BRUCK CLAUDINE ELVIRE MARIE (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ;
            VINALS Y DE BASSOLS CARLOTA (BE) ; COCHE THIERRY (BE) ; CASSART
            JEAN POL (BE)
FEATURES             Location/Qualifiers
     source          1..498
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
Qy    Best Local Similarity   31.9%;
      Matches  475;  Conservative  0;  Score 453;  DB 6;  Length 498;
                                       Best Local Similarity  99.6%;  Pred. No. 3.6e-69;
                                       Matches  475;  Conservative  0;  Mismatches  0;  Indels  2;  Gaps  2;
```

```
Qy    947 CTCTAGCGTCGCCGGCTCGTCGGCCCGGCT-CCGCCCTAGGCTCTCCGCGCTCACCACTCAGTACG 1005
          |||||| || ||| ||| || |||||||| | |||| ||||||||||||||| |||||||||||||
Db      1 CTCTAGGGTGCCGGCTCGTCGGCCCGGCTCCCGCCCTAGGCTCTCCGCGCTCACCACTCAGTACG 60

Qy   1006 -GGGCGCCGCCTTCGGGTCAGCGTCGGCAACCGGCGGGCGGTTCCTCGGAGGGCGGGCC 1064
           ||||| ||||||| |||| |||| || ||||||||| || ||||||||||| |||||||
Db     61 AGGGCGCCGCCTTCTGGGTCGGCGGCCAACCGGCGGCGGTTCCTCGGAGGGCC 120

Qy   1065 GTGGTGAGTCTCCAGTATGTTCGGCCCGCCCGCTCTTGGGACCAAAGCGCC 1124
          ||||||||||| || || |||| |||| |||| || |||| || |||||||
Db    121 GTGGTGAGTCTCCAGTATGTTCGGCCCGCCCGCTCTTGGGACCAAAGCGCC 180

Qy   1125 AAGGACTGCAGCCAGGAGGAGAGGGGGGCTCACCTGCCACCACTGCACAAG 1184
          ||||||||||||||||| ||||| ||| || ||| | |||| |||| ||||
Db    181 AAGGACTGCAGCCAGGAGGAGAGGGGGGCTCACCTCTTATCCTGGGCCGCACCACTGCACAAG 240

Qy   1185 CAGGCCGGCTCTCCCAGACTTAAAATGTATCACCACTAACCTGTGAGGGGGGACCCAAATCTG 1244
          |||| ||||||||||||| ||| || |||  |||||||||| ||| |||||| ||| |||||
Db    241 CAGGCCGGCTCTCCCAGACTTAAAATGTATCACCACTAACCTGTGAGGGGGGACCCAAATCTG 300

Db    301 GACTCCTTCCCCGCGCCTTGGGACATCGCAGGCGCGGGAGAAGCAGTGCCCGCGCCAGGGCC 360

Qy   1305 AGGAGAGCTCCAGGAAGGGGCACTGAGCGGCTCGTGCCGCCGAGGCCTTCGGACATCCGCCAGCC 1364
          ||||||||||||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||
Db    361 AGGAGAGCTCCAGGAAGGGGCACTGAGCGGCTCGTGCCGCCGAGGCCTTCGGACATCCGCCAGGC 420

Qy   1365 ACCAGGGAAAGTCTCCTGGGGGCGGATCTGTAAATAAAACCTTTTTTTTTTTGTTTTTT 1421
          ||||||||||||||||||||| ||||||||||||||||| |||||||||||||||||||
Db    421 ACCAGGGAAAGTCTCCTGGGGGCGGATCTGTAAATAAAACCTTTTTTTTTTTGTTTTTT 477
```

RESULT 7
```
LOCUS       AC012255              166937 bp    DNA     linear   HTG 06-MAY-2001
DEFINITION  Homo sapiens chromosome 15 clone RP11-109D20 map 15, WORKING DRAFT
            SEQUENCE, 22 unordered pieces.
ACCESSION   AC012255
VERSION     AC012255.4  GI:7684440
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 166937)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-109D20
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 166937)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 3, 2000 this sequence version replaced gi:7209928.
            All repeats were identified using RepeatMasker;
            Smit,A.F.A. & Green,P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            --------------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            --------------- Project Information
            Center project name: L3737
            Center clone name: 109_D_20
            --------------- Summary Statistics
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 153800 bases at least Q40
            Consensus quality: 160006 bases at least Q30
            Consensus quality: 163275 bases at least Q20
            Insert size: 164837; sum-of-contigs
            Quality coverage: 4.0 in Q20 bases; sum-of-contigs
            ---------------
```

```
*  NOTE: This is a 'working draft' sequence. It currently
*  consists of 22 contigs. The true order of the pieces
*  is not known and their order in this sequence record is
*  arbitrary. Gaps between the contigs are represented as
*  runs of N, but the exact sizes of the gaps are unknown.
*  This record will be updated with the finished sequence
*  as soon as it is available and the accession number will
*  be preserved.
*
*       1      2087: contig of 2087 bp in length
*    2088      2187: gap of 100 bp
*    2188      3626: contig of 1439 bp in length
*    3627      3726: gap of 100 bp
*    3727      4845: contig of 1119 bp in length
*    4846      4945: gap of 100 bp
*    4946      6702: contig of 1757 bp in length
*    6703      6802: gap of 100 bp
*    6803      8326: contig of 1524 bp in length
*    8327      8426: gap of 100 bp
*    8427     10494: contig of 2068 bp in length
*   10495     10594: gap of 100 bp
*   10595     13215: contig of 2621 bp in length
*   13216     13315: gap of 100 bp
*   13316     16451: contig of 3136 bp in length
*   16452     16551: gap of 100 bp
*   16552     20843: contig of 4292 bp in length
*   20844     20943: gap of 100 bp
*   20944     24940: contig of 3997 bp in length
*   24941     25040: gap of 100 bp
*   25041     29901: contig of 4861 bp in length
*   29902     30001: gap of 100 bp
*   30002     34189: contig of 4188 bp in length
*   34190     34289: gap of 100 bp
*   34290     38386: contig of 4097 bp in length
*   38387     38486: gap of 100 bp
*   38487     45827: contig of 7341 bp in length
*   45828     45927: gap of 100 bp
*   45928     51729: contig of 5802 bp in length
*   51730     51829: gap of 100 bp
*   51830     60123: contig of 8294 bp in length
*   60124     60223: gap of 100 bp
*   60224     73532: contig of 13309 bp in length
*   73533     73632: gap of 100 bp
*   73633     87433: contig of 13801 bp in length
*   87434     87533: gap of 100 bp
*   87534    101708: contig of 14175 bp in length
*  101709    101808: gap of 100 bp
*  101809    121193: contig of 19385 bp in length
*  121194    121293: gap of 100 bp
*  121294    139051: contig of 17758 bp in length
*  139052    139151: gap of 100 bp
*  139152    166937: contig of 27786 bp in length.

FEATURES             Location/Qualifiers
     source          1..166937
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP11-109D20"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1..2087
                     /note="assembly_fragment"
     gap             2088..2187
                     /estimated_length=100
     misc_feature    2188..3626
                     /note="assembly_fragment"
     gap             3627..3726
                     /estimated_length=100
     misc_feature    3727..4845
                     /note="assembly_fragment"
     gap             4846..4945
                     /estimated_length=100
     misc_feature    4946..6702
                     /note="assembly_fragment"
     gap             6703..6802
                     /estimated_length=100
     misc_feature    6803..8326
                     /note="assembly_fragment"
     gap             8327..8426
                     /estimated_length=100
     misc_feature    8427..10494
                     /note="assembly_fragment"
     gap             10495..10594
                     /estimated_length=100
     misc_feature    10595..13215
                     /note="assembly_fragment"
     gap             13216..13315
                     /estimated_length=100
     misc_feature    13316..16451
                     /note="assembly_fragment"
     gap             16452..16551
                     /estimated_length=100
     misc_feature    16552..20843
                     /note="assembly_fragment"
     gap             20844..20943
                     /estimated_length=100
     misc_feature    20944..24940
                     /note="assembly_fragment"
     gap             24941..25040
                     /estimated_length=100
     misc_feature    25041..29901
                     /note="assembly_fragment"
     gap             29902..30001
                     /estimated_length=100
     misc_feature    30002..34189
                     /note="assembly_fragment"
     gap             34190..34289
                     /estimated_length=100
     misc_feature    34290..38386
                     /note="assembly_fragment"
     gap             38387..38486
                     /estimated_length=100
     misc_feature    38487..45827
                     /note="assembly_fragment"
     gap             45828..45927
                     /estimated_length=100
     misc_feature    45928..51729
                     /note="assembly_fragment"
     gap             51730..51829
                     /estimated_length=100
     misc_feature    51830..60123
                     /note="assembly_fragment"
     gap             60124..60223
                     /estimated_length=100
     misc_feature    60224..73532
                     /note="assembly_fragment"
     gap             73533..73632
                     /estimated_length=100
     misc_feature    73633..87433
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     gap             87434..87533
                     /estimated_length=100
     misc_feature    87534..101708
                     /note="assembly_fragment"
     gap             101709..101808
                     /estimated_length=100
     misc_feature    101809..121193
                     /note="assembly_fragment"
     gap             121194..121293
                     /estimated_length=100
     misc_feature    121294..139051
                     /note="assembly_fragment"
     gap             139052..139151
                     /estimated_length=100
     misc_feature    139152..139151
```

RESULT 8

```
Query Match    29.1%;    Score 413.4;   DB 14;   Length 166937;
Best Local Similarity   98.6%;   Pred. No. 1.3e-62;
Matches  417;   Conservative    0;   Mismatches    6;   Indels    0;   Gaps    0;

Qy          1 GAACCAGGAAAAGTAACGGCTACAGACAGTGAGAAATAGTTCGGCTCGGCGGGCTTAGAAAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     157708 GAACCAGGAAAAGTAACGGCTACAGACAGTGAGAAATAGTTCGGCTCGGCGGGCTTAGAAAA 157767

Qy         61 CTCTGTCGTACCAAACCCCAGAGCGGTTGAGAGCGAACAGTGAGAAATAGTTCCTTAACGGA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     157768 CTCTGTCGTACCAAACCCCAGAGCGGTTGAGAGCGAACAGTGAGAAATAGTTCCTTAACGGA 157827

Qy        121 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCGCTTGTACGCAAAGAGACG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     157828 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCGCTTGTACGCAAAGAGACG 157887

Qy        181 CCAAGGACGCGCGCTTCTCCCGCGTCCAGGCAGCCCCAGCCTTGCCTGGCTTGCCCGCCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     157888 CCAAGGACGCGCGCTTCTCCCGCGTCCAGGCAGCCCCAGCCTTGCCTGGCTTGCCCGCCTG 157947

Qy        241 CGTGCAGCACTCCGCGCCGGGCGTGCAGCATGACCCGTGGAGGTACTGCCTTTTTACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     157948 CGTGCAGCACTCCGCGCCGGGCGTGCAGCATGACCCGTGGAGGTACTGCCTTTTTACC 158007

Qy        301 CCCAGCCCCGGCCATGCCCGCCAGGGTTCAGCGGTTCACTCGCTCATCGGTTATTCTAGTGTTTT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     158008 CCCAGCCCCGGCCATGCCCGCCAGGGTTCAGCGGTTCACTCGCTCATCGGTTATTCTAGTGTTTT 158067

Qy        361 TGGCTCTAGCAGCAAGCTTCGTGTCATCTTGGCCGGGGATCCGGTCCACTCGCGGGGCTCGGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     158068 TGGCTCTAGCAGCAAGCTTCCTGTCTCATCTTGCGCGGGGATCCGGTGGCCACTCGGTAAGGG 158127

Qy        421 TTT 423
              ||
Db     158128 TGT 158130
```

```
LOCUS       AC091117            181312 bp    DNA     linear   PRI 20-FEB-2002
DEFINITION  Homo sapiens chromosome 15 clone RP11-109D20 map 15q15, complete
            sequence.
ACCESSION   AC091117
VERSION     AC091117.5  GI:18767395
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 181312)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 181312)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2001) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
REFERENCE   3  (bases 1 to 181312)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-2002) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
REFERENCE   4  (bases 1 to 181312)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-2002) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
REFERENCE   5  (bases 1 to 181312)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2002) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
COMMENT     On Feb 20, 2002 this sequence version replaced gi:18653547.
            -------------------- Genome Center
            Center: Multimegabase Sequencing Center
            Center code: UWMSC
            Web site: http://chroma.mbt.washington.edu/msg_www
            Contact: leerowen@systemsbiology.org
            -------------------- Summary Statistics
            Sequencing vector: pUC18; L08752
            Chemistry: Dye-terminator Big Dye; 90% of reads
            Chemistry: Dye-primer Big Dye; 10% of reads
            Assembly program: Phrap; version 0.990399

            Note: Data from overlapping clone AC090888 [Drafting center: WIBR]
            was added for finishing. RP11-109D20 spans bases 1-174236 of this
            sequence. Bases 174237-179373 derive from clone RP11-276K9,
            AC087790 [Drafting center: WIBR] to establish the overlap with the
            next BAC in our tiling path. Bases 179373-181312 derive from
            finished clone CTD-2651B20, AC051619 [Drafting center: UWMSC] in
            order to give overlap required for determining long range
            contiguity.
FEATURES             Location/Qualifiers
     source          1..181312
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q15"
                     /clone="RP11-109D20"
                     /clone_lib="RPCI human BAC library 11"
                     /note="This clone overlaps CTD-2014N11 AC090888 and
                     RP11-276K9 AC087790. Data from overlapping BACs were added
                     and the consensus sequence determined from RP11-109D20 to
                     the extent possible."
     misc_feature    1..8913?
                     /note="overlap with CTD-2014N11, AC090888"
     unsure          36107..36120
                     /note="low quality data"
     misc_feature    40470..40820
                     /note="sequence data generated from subcloned PCR product"
     misc_feature    53030..53480
                     /note="sequence data generated from subcloned PCR product"
     unsure          86982..86983
                     /note="low quality data"
     misc_feature    95610..96030
                     /note="sequence data generated from subcloned PCR product"
     misc_feature    96830..97420
                     /note="sequence data generated from subcloned PCR product"
     misc_feature    98260..98420
                     /note="single clone coverage"
     unsure          98850..98852
                     /note="low quality data"
     unsure          10981..101003
                     /note="low quality data"
     misc_feature    174237..179372
                     /note="overlap with RP11-276K9, AC087790. This sequence
                     data was added to establish the overlap between the BACs
                     in the tiling path"
```

```
misc_feature    179373..181312
                /note="overlap with CTD-2651B20, AC051619. This data is
                added to give overlap for long range contiguity."
ORIGIN

Query Match     29.1%;   Score 413.4;  DB 8;  Length 181312;
Best Local Similarity 98.6%;  Pred. No. 1.2e-62;
Matches 417;  Conservative 0;  Mismatches 6;  Indels 0;  Gaps 0;

Qy      1 GAACCAGGAAAGTAAGGCTACACACAGTGAGAAATAGTTTCGCTCGGCGCCTAGAAAAA     60
          ||||||||||||||||| |||||| |||||||||||||||||||||||||||||||||||
Db 165008 GAACCAGGAAAGTAAGCGCTACACAGACAGTGAGAAATAGTTTCGCTCGGCGCCTAGAAAAA 165067

Qy     61 CTCTGTCGGTACCAACCCAGAGCGGGCTTGAGAGCAGCCAGCCACCTCCACGCTTCCTTAACGGA    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 165068 CTCTGTCGGTACCAACCCAGAGCGGGCTTGAGAGCAGCCAGCCACCTCCACGCTTCCTTAACGGA 165127

Qy    121 GAGGTGCAGGACTCACCAGACTTCACCAGCCCACTCGGTCCCAGCCGCCTTGTACGGAAAGAGACG    180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 165128 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCCCTTGTACGGCCAAAGAGACG 165187

Qy    181 CCAAGGACGCGCGCTCTCCCCCGGTCCAGGCAGCCCAGCGCGTTGCTGGCTTGCCTGCCCGGCCTG    240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 165188 CCAAGGACGCGCGCTCTCCCCCGGTCCAGGCAGCCCAGCTTCCTGGCTTGCCTGCCCGGCCTG 165247

Qy    241 CGTGCGACACTCGGCGCGGGTGCAGCATGACCCTGTGGAACGGCGTACTGCCTTTTTTACC    300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 165248 CGTGCGACACTCGGCGCGTGCGACCATGACCCTGTGGAACGGCGTACTGCCTTTTTTACC 165307

Qy    301 CCCAGCCCCGGCCATGCCCCGGCCTTCAGCGTTCCACTGCTCCATCGTTATTCTAGTGTTTT    360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 165308 CCCAGCCCCGGCCATGCCGCGAGGCTTCAGCGTTCCACTGCTCCATCGTTATTCTAGTGTTTT 165367

Qy    361 TGGGCTTAGCGAGCAAGCTTCCTGCTCGTCATCTTGCCCGGGGAGATCCGTCGCCCACTCGGCGCTCGGT    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 165368 TGGGCTTAGCGAGCAAGCTTCCTGCTCGTCATCTTGCCCGGGGAATCCGTCGCCACTCGGTAAGGG 165427

Qy    421 TTT 423
          | |
Db 165428 TGT 165430
```

RESULT 9

```
LOCUS       AC009700/c              156534 bp    DNA     linear   HTG 20-APR-2000
DEFINITION  Homo sapiens chromosome 15 clone RP11-163P10 map 15, WORKING DRAFT
            SEQUENCE, 15 unordered pieces.
ACCESSION   AC009700
VERSION     AC009700.4  GI:7622346
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 156534)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 156534)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
            Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Depayre,E., Devon,K., Dewar,K.,
            Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
            Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
            Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
            Karatas,A., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P.,
            Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
            Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Naylor,J., Nioff,M., O'Connor,T., O'Donnell,P., Pavlin,D.,
            Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
            Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
            Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 20, 2000 this sequence version replaced gi:6056272.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
            Center project name: L2336
            Center clone name: 163_P_10
            -------------- Summary Statistics
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 108837 bases at least Q40
            Consensus quality: 130230 bases at least Q30
            Consensus quality: 145227 bases at least Q20
            Insert size: 157000; agarose-fp
            Insert size: 155134; sum-of-contigs
            Quality coverage: 3.2 in Q20 bases; agarose-fp
            Quality coverage: 3.3 in Q20 bases; sum-of-contigs
            --------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 15 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            *       1     1169: contig of 1169 bp in length
            *    1170     1269: gap of 100 bp
            *    1270     3335: contig of 2066 bp in length
            *    3336     3435: gap of 100 bp
            *    3436     5698: contig of 2263 bp in length
            *    5699     5798: gap of 100 bp
            *    5799     9216: contig of 3418 bp in length
            *    9217     9316: gap of 100 bp
            *    9317    15633: contig of 6317 bp in length
            *   15634    15733: gap of 100 bp
            *   15734    21746: contig of 6013 bp in length
            *   21747    21846: gap of 100 bp
            *   21847    30351: contig of 8505 bp in length
            *   30352    30451: gap of 100 bp
            *   30452    36509: contig of 6058 bp in length
            *   36510    36609: gap of 100 bp
            *   36610    45279: contig of 8670 bp in length
            *   45280    45379: gap of 100 bp
            *   45380    56851: contig of 11472 bp in length
            *   56852    56951: gap of 100 bp
            *   56952    67374: contig of 10423 bp in length
            *   67375    67474: gap of 100 bp
            *   67475    79534: contig of 12060 bp in length
            *   79535    79634: gap of 100 bp
            *   79635    95432: contig of 15798 bp in length
            *   95433    95532: gap of 100 bp
            *   95533   120801: contig of 25269 bp in length
            *  120802   120901: gap of 100 bp
            *  120902   156534: contig of 35633 bp in length
FEATURES             Location/Qualifiers
     source          1..156534
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
```

```
misc_feature   1..1169
               /clone="RP11-163P10"
               /clone_lib="RPCI-11 Human Male BAC"
               /note="assembly_fragment"
gap            1170..1269
               /estimated_length=100
misc_feature   1270..3335
               /note="assembly_fragment"
gap            3336..3435
               /estimated_length=100
misc_feature   3436..5698
               /note="assembly_fragment"
gap            5699..5798
               /estimated_length=100
misc_feature   5799..9216
               /note="assembly_fragment"
gap            9217..9316
               /estimated_length=100
misc_feature   9317..15633
               /note="assembly_fragment"
gap            15634..15733
               /estimated_length=100
misc_feature   15734..21746
               /note="assembly_fragment"
gap            21747..21846
               /estimated_length=100
misc_feature   21847..30351
               /note="assembly_fragment"
gap            30352..30451
               /estimated_length=100
misc_feature   30452..36509
               /note="assembly_fragment"
               /clone_end:T7
               vector_side:left"
gap            36510..36609
               /estimated_length=100
misc_feature   36610..45279
               /note="assembly_fragment"
gap            45280..45379
               /estimated_length=100
misc_feature   45380..56851
               /note="assembly_fragment"
gap            56852..56951
               /estimated_length=100
misc_feature   56952..67374
               /note="assembly_fragment"
gap            67375..67474
               /estimated_length=100
misc_feature   67475..79534
               /note="assembly_fragment"
gap            79535..79634
               /estimated_length=100
misc_feature   79635..95432
               /note="assembly_fragment"
gap            95433..95532
               /estimated_length=100
misc_feature   95533..120801
               /note="assembly_fragment"
gap            120802..120901
               /estimated_length=100
misc_feature   120902..156534
               /note="assembly_fragment"

ORIGIN

Query Match    29.0%;  Score 411.8;  DB 14;  Length 156534;
Best Local Similarity   98.3%;  Pred. No. 2.4e-62;
Matches  416;  Conservative  0;  Mismatches  7;  Indels  0;  Gaps  0;

Qy    1 GAACCAGAAAAGTAAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCCGCCGCTTAGAAAAA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7156 GAACCAGAAAAGTAAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCCGCCGCTTAGAAAAA 7097

Qy   61 CTCTGTCGGTACCAAACCCAGAGCGTTGAGAGCAGCCCACCTCCACGCGCTTCCTTAACGGA 120
```

```
RESULT 10
LOCUS       AC087790      171444 bp    DNA    linear    HTG 09-MAY-2001
DEFINITION  Homo sapiens chromosome 15 clone RP11-276K9 map 15, WORKING DRAFT
            SEQUENCE, 13 unordered pieces.
ACCESSION   AC087790
VERSION     AC087790.2  GI:13357344
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 171444)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-276K9
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 171444)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,V., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
            Gardyna,S., Ginde,S., Goyette,M., Graham,M., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iley,I., Johnson,R.,
            Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
            Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
            Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
            Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
            Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,B., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Travis,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JAN-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 16, 2001 this sequence version replaced gi:12408501.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
```

```
Db 7096 CTCTGTCGGTACCAAACCCAGAGCGTTGAGAGCAGCCCACCTCCACGCGCTTCCTTAACGGA 7037

Qy  121 GAGGTGCAGGACTTCACGAGACTTCACCAGCCCATCGGTCCTCCAGCCTTGTACGGCAAAGAGACG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7036 GAGGTGCAGGACTTCACGAGACTTCACCAGCCCATCGGTCCTCCAGCCTTGTACGGCAAAGAGACG 6977

Qy  181 CCAAGGACGCGGCGCTCTCCCGCGGTCCAGGCAGCCCACTCGGTCCCAGCCGTTCCCGGCCTG 240
         ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6976 TCAAGGACGCGGCGCTCTCCCGCGGTCCAGGCAGCCCACTCGGTCCCAGCCGTTCCCGGCCTG 6917

Qy  241 CGTGCAGCACTCGGCGCCGGCGCGTGCAGCATGACCCGTGTGGGAACGGGGGTACTGCCTTTACC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6916 CGTGCAGCACTCGGCGCCGGCGCGTGCAGCATGACCCGTGTGGGAACGGGGGTACTGCCTTTTACC 6857

Qy  301 CCCAGCCCCGGCCATGCCGCAGGCGTTCAGCGGTTCCACTGTTGTGATCTGTTATTCTAGTTGTTT 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6856 CCCAGCCCCGGCCATGCCGCAGGCGTTCAGCGGTTCCACTGTTGTGATCTGTTATTCTAGTTGTTT 6797

Qy  361 TGGGCTCTAGCAGCAAGCTTCCTGCTCATCTTGCCGGGGGATCCGGTGGGCCACTCGCGCTGGT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6796 TGGGCTCTAGCAGCAAGCTTCCTGCTCATCTTGCCGGGGGATCCGGTGGGCCACTCGGTGGG 6737

Qy  421 TTT 423
         | |
Db 6736 TGT 6734
```

http://ftp.genome.washington.edu/RM/RepeatMasker.html

```
-------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-------- Project Information
Center project name: L12451
Center clone name: 276_K_9
-------- Summary Statistics
Sequencing Vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 165906 bases at least Q40
Consensus quality: 168587 bases at least Q30
Consensus quality: 169557 bases at least Q20
Insert size: 180000; agarose-fp
Insert size: 170244; sum-of-contigs
Quality coverage: 5.7 in Q20 bases; agarose-fp
Quality coverage: 6.0 in Q20 bases; sum-of-contigs
```

```
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

```
*        1   6091: contig of 6091 bp in length
*     6092   6191: gap of 100 bp
*     6192   7986: contig of 1795 bp in length
*     7987   8086: gap of 100 bp
*     8087  43939: contig of 35853 bp in length
*    43940  44039: gap of 100 bp
*    44040  45298: contig of 1259 bp in length
*    45299  45398: gap of 100 bp
*    45399  47595: contig of 2197 bp in length
*    47596  47695: gap of 100 bp
*    47696  52385: contig of 4690 bp in length
*    52386  52485: gap of 100 bp
*    52486  56310: contig of 3825 bp in length
*    56311  56410: gap of 100 bp
*    56411  71313: contig of 14903 bp in length
*    71314  71413: gap of 100 bp
*    71414  90993: contig of 19580 bp in length
*    90994  91093: gap of 100 bp
*    91094 111459: contig of 20366 bp in length
*   111460 111559: gap of 100 bp
*   111560 137726: contig of 26167 bp in length
*   137727 137826: gap of 100 bp
*   137827 170053: contig of 32227 bp in length
*   170054 170153: gap of 100 bp
*   170154 171444: contig of 1291 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..171444
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP11-276K9"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1..6091
                     /note="assembly_fragment
                     clone:SP6
                     vector_side:left"
     gap             6092..6191
                     /estimated_length=100
     misc_feature    6192..7986
                     /note="assembly_fragment"
     gap             7987..8086
                     /estimated_length=100
```

```
     misc_feature    8087..43939
                     /note="assembly_fragment"
     gap             43940..44039
                     /estimated_length=100
     misc_feature    44040..45298
                     /note="assembly_fragment"
     gap             45299..45398
                     /estimated_length=100
     misc_feature    45399..47595
                     /note="assembly_fragment"
     gap             47596..47695
                     /estimated_length=100
     misc_feature    47696..52385
                     /note="assembly_fragment"
     gap             52386..52485
                     /estimated_length=100
     misc_feature    52486..56310
                     /note="assembly_fragment"
     gap             56311..56410
                     /estimated_length=100
     misc_feature    56411..71313
                     /note="assembly_fragment"
     gap             71314..71413
                     /estimated_length=100
     misc_feature    71414..90993
                     /note="assembly_fragment"
     gap             90994..91093
                     /estimated_length=100
     misc_feature    91094..111459
                     /note="assembly_fragment"
     gap             111460..111559
                     /estimated_length=100
     misc_feature    111560..137726
                     /note="assembly_fragment"
     gap             137727..137826
                     /estimated_length=100
     misc_feature    137827..170053
                     /note="assembly_fragment"
     gap             170054..170153
                     /estimated_length=100
     misc_feature    170154..171444
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
```

```
ORIGIN

Query Match      29.0%;  Score 411.8; DB 14;  Length 171444;
Best Local Similarity   98.3%;  Pred. No. 2.4e-62;
Matches 416;  Conservative   0;  Mismatches   7;  Indels   0;  Gaps   0;

Qy        1 GAACCAGGAAAGTAACCGGCTACAGACAGTGAGAAATAGTTTGCTCGCGGCGCTAGAAAAA  60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140405 GAACCAGGAAAGTAACCGGCTACAGACAGTGAGAAATAGTTTGCTCGCGGCGCTAGAAAAA 140464

Qy       61 CTCTGTCGGGTACCAACCCCAGAGCGTTGAGGAGCAGCCCACCTCCACCGCTTCCTTAACGGA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140465 CTCTGTCGGGTACCAACCCCAGAGCGTTGAGGAGCAGCCCACCTCCACCGCTTCCTTAACGGA 140524

Qy      121 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCCTTGTACGCAAAGAGACG 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140525 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCCTTGTACGCAAAGAGACG 140584

Qy      181 CCAAGGACGCGCTCTCCCGCGTCCAGGCGAGCCCCAGCTTGCTGGCTTGCCTGGCCGCTG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140585 TCAAGGACGCGCTCTCCCGCGTCCAGGCAGCCCCAGCGTTGCTGGCTTGCCTGGCCGCTG 140644

Qy      241 CGTGCAGCACTCGGGCGCCGGCGTCGCAGCAATGACCCCTGTGGAACGGCGTACTGCCTG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140645 CGTGCAGCACTCGGGCGCCGGCGTCGCAGCAATGACCCCTGTGGAACGGCGTACTGCCTG 140704

Qy      301 CCCAGCCCCGGCATGCCGCAGCCTTCAGGCGTTCCACTGCTCATCGTTAATTCTTAATTT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140705 CCCAGCCCCGGCATGCCGCAGCCTTCAGGCGTTCCACTGCTCATCGTTATTCTTAGTGTTT 140764
```

RESULT 11

LOCUS   AX714511   2684 bp   DNA   linear   PAT 15-APR-2003
DEFINITION Sequence 1195 from Patent EP1293569.
ACCESSION   AX714511
VERSION   AX714511.1 GI:29894464
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
            Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
            Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
            Masuho,Y.
  TITLE   Full-length cDNAs
  JOURNAL   Patent: EP 1293569-A 1195 19-MAR-2003;
            Helix Research Institute (JP) ; Research Association for
            Biotechnology (JP)
FEATURES   Location/Qualifiers
     source   1..2684
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   27.5%; Score 391; DB 6; Length 2684;
Best Local Similarity   97.5%; Pred. No. 1.9e-58;
Matches   397; Conservative   0; Mismatches 10; Indels   0; Gaps   0;

Qy   361 TGGCTCTPAGCAGCAAGCTTCCTGCTCATCTTGCCGGGGGATCCGGTTGGGCCACTCGCGCGCTGGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140765 TGGCTCTPAGCAGCAAGCTTCCTGCTCATCTTGCCGGGGGATCCGGTTGGGCCACTCGGTAAGGG 140824

Qy   421 TTT 423
              |||
Db   140825 TGT 140827

RESULT 12

LOCUS   AK056896   2684 bp   mRNA   linear   PRI 30-JAN-2004
DEFINITION Homo sapiens cDNA FLJ32334 fis, clone PROST2005426.
ACCESSION   AK056896
VERSION   AK056896.1 GI:16552419
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
            Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
            Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
            Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
            Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
            Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
            Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
            Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
            Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
            Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
            Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
            Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
            Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
            Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
            Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
            Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
  TITLE   Complete sequencing and characterization of 21,243 full-length
            human cDNAs
  JOURNAL   Nat. Genet. 36 (1), 40-45 (2004)
   PUBMED   14702039
REFERENCE   2
  AUTHORS   Isogai,T., Otsuki,T. and Sugiyama,T.
  TITLE   Direct Submission
  JOURNAL   Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: RAB and
            HRI.
FEATURES   Location/Qualifiers
     source   1..2684
                     /organism="Homo sapiens"

Qy   1015 TCTGGGTCACGCTGGCAACCGGGCGTGGTCCCTGTGCCTCTTCCTCGGAGGGGCGTGGTGAGTC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2363 TTTCGAGTCCCACCGCCACAGGGTCCGTGGCCTCCTTCCTTCCTCGGAGGGCGTGGTGAGTC 2304

Qy   1075 TCCAGTAAGTTCGGCCGCCCAGCGCTGCTTCTTCGACCAAAGGCGCCAAGGACTGCA 1134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2303 TCCAGTPATGTTCGGCCCAGCGCTCCTTCGCACCCTTCTTCGACCAAAGCGCCAAAGGACTGCA 2244

Qy   1135 GCCAGGAGGAGGAGGGGGGCTCACCTCTTATCTCGGCGACAAGCAGCAGGCGCGCTC 1194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2243 GCCAGGAGGAGGAGGGGGGCTCACCTCTTATCCTCGGCGACAAGCAGCAGGCGCGCTC 2184

Qy   1195 TCCCAGACTTAAAATGTATCACCACTAACCACTAACCTGTGAGGGGGACCCAAATCTGGACTCCTTCC 1254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2183 TCCCAGACTTAAAATGTATCACCACTAACCTGTGAGGGGGACCCAAATCTGGACTCCTTCC 2124

Qy   1255 CCGGCCTTGGGACATCCGCAGGCGGGAAGCAGTGCCCGCCAGGGCTGGGCCAGGAGAGCTC 1314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2123 CCGGCCTTGGGACATCGGCAGGCGGGGAAGCAGTGCCCGCCAGGGCTGGGCCAGGAGAGCTC 2064

Qy   1315 CAGGAAGGGCACTGAGCGCTGCTGGGGCAGCACTCCCAGGGGACATCCGCCAGGGACACCAGGGGAA 1374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2063 CAGGAAGGGCACTGAGCGCTGCTGGGGCAGCACTCCCAGGGGACATCCGCAGGGACACCAGGGGAA 2004

Qy   1375 GTCTCCTGGGGGGGGATCTGTAAATAAACCTTTTTTTCTTTGTTTTTTT 1421
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2003 GTCTCCTGCGGGGGGGATCTGTAAATAAACCTTTTTTTTCTTTGTTTTTTT 1957

```
CDS             56. .952
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="PROST2005426"
                /tissue_type="prostate"
                /clone_lib="PROST2"
                /note="cloning vector: pME18SFl3"
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="BAB71304.1"
                /db_xref="GI:16552420"
                /translation="MATLGHTPFYAGPKPTPFMDTTLASIIMIFLTALATFIVILPG
                IRGKTRLFWLLRVVSLPIGAAILGTPVQQLNETINYNEBFTWRLGENYABEYAKALE
                KGLPDVLYLAEKFTPRSPCGLYRLAGHYTSAMLWVAFLCWLLANVMLSMPVLVY
                GGYMLLATGIFQLLALLFFSMATSLTSPCPLHLGASVLHTHHGPAFWITLLTTGLLCVL
                LGLAWAVAHRMQPHRLKAFPNQSVDEDPMLEWSPEGGLLSPRYRSMADSPKSQDIFPL
                SBASSTVAYCKEAHPKDPDCAL"

ORIGIN
Best Local Similarity  97.5%;   Pred. No. 1.9e-58;
Query Match           27.5%;   Score 391;  DB 8; Length 2684;
Matches 397; Conservative   0; Mismatches 10; Indels  0; Gaps  0;

Qy    1015 TCTGGGTCACGCTGGCACCGGCGGGTGCTCCGTGCCTCTTCCTCGCGAGGGGCGGTGCTGTGAGTC 1074
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2363 TTTCGATCCCCACCGCCACAGGCGTCCTGTGCCTCTTCCTCGCGAGGGGCCGTGGTGAGTC 2304

Qy    1075 TCCAGTATGTTCCGCCCAGCGCTCTTCGCACCCTTCTCGCACCCTTCCTCGGACCAAAAGCCCCAAGGACTGCA 1134
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2303 TCCAGTATGTTCCGCCCAGCGCTCTTCGCACCCTTCTCGACCAAAAGCGCCAAGGACTGCA 2244

Qy    1135 GCCAGGAGGAGGGGGCGCTCACCTTCTTATCCTCGGGCGACCCCACTGCACAAGCAGGCCGCGCTC 1194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2243 GCCAGGAGGAGGGGGCTCACCTTCTTATCCTCGGGCGACCCCACTGCACAAGCAGGCCGCGCTC 2184

Qy    1195 TCCCAGACTTAAAATGTATCACCACTAACCTGTGAGGGGGACCCAATCTGGACTCCTTCC 1254
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1195 TCCCAGACTTAAAATGTATCACCACTAACCTGTGAGGGGGACCCAATCTGGACTCCTTCC 2124

Qy    1255 CCGCCTGGGACATCGCCAGGCCGGGAAGCAGTGCCCGCCGCCGCTGGGCCCCAGGAGAGCTC 1314
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2123 CCGCCTGGGACATCGCCAGGCCGGGAAGCAGTGCCCGCCGCCGCTGGGCCCCAGGAGAGCTC 2064

Qy    1315 CAGGAAGGGCACTGAGCGCGCTGCTGGCGCGCGGGACATCCGCAGGCCACCAGGGAAA 1374
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2063 CAGGAAGGGCACTGAGCGCGCTGCTGGCGCGCGGGACATCCGCAGGCCACCAGGGAAA 2004

Qy    1375 GTCTCCTGGGGCGGATCTGTAAATAAAACCTTTTTTTTTTTTTTTGTTTTTT 1421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2003 GTCTCCTGGGGCGGATCTGTAAATAAAACCTTTTTTTTTTTTTTTGTTTTTT 1957
```

RESULT 13

```
*      1      1169: contig of 1169 bp in length
*   1170      1269: gap of 100 bp
*   1270      3335: contig of 2066 bp in length
*   3336      3435: gap of 100 bp
*   3436      5698: contig of 2263 bp in length
*   5699      5798: gap of 100 bp
*   5799      9216: contig of 3418 bp in length
*   9217      9316: gap of 100 bp
*   9317     15633: contig of 6317 bp in length
*  15634     15733: gap of 100 bp
*  15734     21746: contig of 6013 bp in length
*  21747     21846: gap of 100 bp
*  21847     30351: contig of 8505 bp in length
*  30352     30451: gap of 100 bp
*  30452     36509: contig of 6058 bp in length
*  36510     36609: gap of 100 bp
*  36610     45279: contig of 8670 bp in length
*  45280     45379: gap of 100 bp
*  45380     56851: contig of 11472 bp in length
*  56852     56951: gap of 100 bp
*  56952     67374: contig of 10423 bp in length
*  67375     67474: gap of 100 bp
*  67475     79534: contig of 12060 bp in length
*  79535     79634: gap of 100 bp
*  79635     95432: contig of 15798 bp in length
*  95433     95532: gap of 100 bp
```

Cooke,P., DeArellano,K., Depayre,B., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,A., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,P., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Tesfaye,S., Tornuella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Smit,A.F.A. & Green,P. (1996-1997)

```
        *    9553    120801: contig of 25269 bp in length
        *   120802   120901: gap of 100 bp
        *   120902   156534: contig of 35633 bp in length.
FEATURES            Location/Qualifiers
     source         1..156534
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /chromosome="15"
                    /map="15"
                    /clone="RP11-163P10"
                    /clone_lib="RPCI-11 Human Male BAC"
     misc_feature   1..1169
                    /note="assembly_fragment"
     gap            1170..1269
                    /estimated_length=100
     misc_feature   1270..3335
                    /note="assembly_fragment"
     gap            3336..3435
                    /estimated_length=100
     misc_feature   3436..5698
                    /note="assembly_fragment"
     gap            5699..5798
                    /estimated_length=100
     misc_feature   5799..9216
                    /note="assembly_fragment"
     gap            9217..9316
                    /estimated_length=100
     misc_feature   9317..15633
                    /note="assembly_fragment"
     gap            15634..15733
                    /estimated_length=100
     misc_feature   15734..21746
                    /note="assembly_fragment"
     gap            21747..21846
                    /estimated_length=100
     misc_feature   21847..30351
                    /note="assembly_fragment"
     gap            30352..30451
                    /estimated_length=100
     misc_feature   30452..36509
                    /note="assembly_fragment"
     gap            36510..36609
                    /estimated_length=100
                    /note="assembly_fragment
                    vector_side:left"
     misc_feature   36610..45279
                    /note="assembly_fragment"
     gap            45280..45379
                    /estimated_length=100
     misc_feature   45380..56851
                    /note="assembly_fragment"
     gap            56852..56951
                    /estimated_length=100
     misc_feature   56952..67374
                    /note="assembly_fragment"
     gap            67375..67474
                    /estimated_length=100
     misc_feature   67475..79534
                    /note="assembly_fragment"
     gap            79535..79634
                    /estimated_length=100
     misc_feature   79635..95432
                    /note="assembly_fragment"
     gap            95433..95532
                    /estimated_length=100
     misc_feature   95533..120801
                    /note="assembly_fragment"
     gap            120802..120901
                    /estimated_length=100
     misc_feature   120902..156534
                    /note="assembly_fragment"
ORIGIN
```

```
Query Match                    27.5%;  Score 391; DB 14;  Length 156534;
Best Local Similarity   97.5%;  Pred. No. 1e-58;
Matches  397;  Conservative    0;  Mismatches  10;  Indels    0;  Gaps    0;

Qy   1015 TCTGGGGTCACGGCTGGCAACCGGCGGGCGGTCCTGTGCCTCTTCCTCGGAGGGCGTGGTGAGTC 1074
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  96111 TTTCGATCCCCACCGCACCAGGGCGTCCTGTGCCTCTTCCTCGGAGGGCGTGGTGAGTC 96170

Qy   1075 TCCAGTATGTTCGGCCCACCGGCGTTCTTCGCACCCTTCTTCTGGACCAAAGGCGCCAAGGACTGCA 1134
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  96171 TCCAGTATGTTCGGCCCACAGGCGGCTTCTTCCGCACCCTTCTTCTGGACCAAAGGCGCCAAGGACTGCA 96230

Qy   1135 GCCAGGAGGAGAGAGGGGGGCTTCACCTCTTATCCTCTGGGCGACCCACTGCACAAAGCACGCCGCTC 1194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  96231 GCCAGGAGGAGAGGGGGGCTTCACCTCTTATCCTCGGGCGACCCACTGCACAAAGCACGCCGCTC 96290

Qy   1195 TCCCAGACTTAAAATGTATTCACCACTAACCTGTTGGAGGGGGGACCCCAATCTTGGACTCCTTCC 1254
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  96291 TCCCAGACTTAAAATGTATTCACCACTAACCTGTTGGAGGGGGGACCCCAATCTTGGACTCCTTCC 96350

Qy   1255 CCCGCCTTCGGACATCGCACAGGGCGCGGGGAAGCAGAGTGTGCCCGGCGGCCTGGGCGTCGGCAGCGAGAGCTC 1314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  96351 CCGCCTTGGGACATCGCAGGGCGCGGACAGCGCCGAAGCAGAGTGGCCCGGCGGCCTGGGCGTCGGCAGCGAGAGCTC 96410

Qy   1315 CAGGAAGGCACTGAGCGCGTGCTGGCGCGGGAGAGCAGAGTGCCCGGCGGCCTGGGCGTCGGCAGGGGGAAA 1374
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  96411 CAGGAAGGCACTGAGCGCGTGCTGGCGGGGAGCAGGCCTGCGGACATCCGCAGGCACCAGGGGGAAA 96470

Qy   1375 GTCTCCCTGGGGGCGGATCCTGTAAATAAACCTTTTTTTTCTTTTTGTTTTTTTTT 1421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  96471 GTCTCCCTGGGGGCGGATCCTGTAAATAAACCTTTTTTTTTTCTTTTTGTTTTTTTTT 96517
```

```
RESULT 14
AX405919
LOCUS        AX405919                 1636 bp     DNA         linear    PAT 14-JUN-2002
DEFINITION   Sequence 334 from Patent WO0222660.
ACCESSION    AX405919
VERSION      AX405919.1  GI:21439277
KEYWORDS     .
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominidae; Homo.
REFERENCE    1
  AUTHORS    Tang,Y.T., Liu,C., Zhou,P., Asundi,V., Zhang,J., Zhao,Q.A., Ren,F.,
             Xue,A.J., Yang,Y., Wehrman,T. and Drmanac,R.T.
  TITLE      Novel nucleic acids and polypeptides
  JOURNAL    Patent: WO 0222660-A 334 21-MAR-2002;
             HYSEQ, INC. (US)
FEATURES            Location/Qualifiers
     source         1..1636
                    /organism="Homo sapiens"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"
     CDS            441..1472
                    /note="unnamed protein product"
                    /codon_start=1
                    /protein_id="CAD35026.1"
                    /db_xref="GI:21439278"
                    /translation="MATLGHTFPYAGPKPTFPMDTTLASIIMLFTLALATFIVILPG
                    IRGKTRLFWLLRVVTSLFIGAAILAVNFSSEWSVGQVSTNTSYXAFSSEWISADIGLQ
                    VGLGGVNITLTGFPVQQLNETINYNEBFTWRLGENYAEEYAKALBKGLPDPVLYLTEK
                    FTPRSPCGLYRQYRLAGHYTSAMLWVAFLCWLLANVMLSMPALVYGGYMLLAGIFQL
                    LALLFFSMATSLISPCPLHLGASVLHTHHGPAFWITLTTGLLCVLLGILAMAVAHRMQP
                    HRLKAFRNQSVDEDPMLEWSPEEGGLLSPRYRSMADSPKSQDIPLSEASSTKAYCKBA
                    HPKDPDCAL"
ORIGIN
```

```
Query Match                    25.1%;  Score 356.4; DB 6;  Length 1636;
Best Local Similarity   63.9%;  Pred. No. 2.3e-52;
```

```
Matches 556;  Conservative  0;  Mismatches 311;  Indels  3;  Gaps  1;

Qy   263 CAGCATGACCCTGTGGAAGGGCGGTACTCGCTGGCTTTTACCCCCAGCCCCGGCATGCGCAGG 322
         |||  ||| |       |              |   |||||       ||| ||
Db   437 CAAGATGGCTACTTTGGGACACACATTCCCTTCTATGCTGGCCCCAGCAGCCAACCTTCCC 496

Qy   323 CTTCAGCGGTTCCACTGCTCATCGTTATTCTACAGTGTTTTTTGGGCTCTAGCAGCTTCCT 382
         |||                    |       ||                |||   ||
Db   497 GATGGACAACCACTTTGGCCAGCATCATCATGACTCTTCACTGCAGTCCACTGGCTTCAT 556

Qy   383 GCTCATCTTGCCGGGGATCCGGTGGGCCACTCGGCGTGGGGTTTTGGTTGGTTGGAGAGTTCTTCT 442
         |||||| ||  |||| |    ||| ||||| |||||                    |||   ||
Db   557 CGTCATCCTGCCCTGGCTTGGGGAAAAGACGAGGCGTTCCTGCTGCTTCGCTGCTGTGAC 616

Qy   443 CAGTCTGTTCATAGGCGCGAGAAATTGTGCTGTCCACTTCAGTGCAGAATGGTTCGTGGG 502
         |||  | || |||| ||   |   |||| ||||  ||||||||| ||             |||
Db   617 CAGCTTATTCATCGGGGCTGCAATCCTGGTGAATTCAGTTCAGGTGGTTCTGGTCGTGGG 676

Qy   503 TACAGTGAAACACCAACACATCCTTACAACAAAGCCTTCAGCGCCGCGTTACAGCCCCGTGT 562
         |||      ||    ||                  |||||||                  |||
Db   677 CCAGGTCAGCCACCAACACAGTTCAGTTCTGAGTGGGATCAGCGCGTGATAT 736

Qy   563 CGGTCTGCTCGTCGTGGGTGGAGGGCCATTAATATTACACAGGGACCCAGTGCATCA 622
         |||||| | ||  ||||| |     ||                     |||||||||
Db   737 TGGGGCTGCAGGTCGGGTGGGGAGTGCAACATCACACTCACAGGGACCCCGTGCAGCA 796

Qy   623 GCTGAACGGAGCAGCATTGACTACAACGACGCAGTGCACTTCACCTGGCGTCTGAAAGAGAATTACGC 682
         |||||| | |||  |          |       ||  ||||   |            |      ||
Db   797 GCTGAATGAGACCATCAATTCACNACGAGAAGTTCACCTGGGTGAGGAACTATGC 856

Qy   683 CGCGGAGTACGGAACGCACTGGGGAGAAGGGCTGCCCGGGAACCCAGTGCTCTCACCTGGGGA 742
         ||||||||| | |||||              |||||| |||||||||| ||| |||| |||||
Db   857 TGAGGAGTATGCAAAGGCTCTGGGGGCTGCCCAGAGCCCTGGTGCCCTACCTCCTGGCTA 916

Qy   743 GAAGTTCACACGAGTAGCCCTGTACCACAGTACCACCTGGCCGGCGGGGACACTA 802
         |||||||||| || ||   |||||||||||||||||||||||||  |||||| |||
Db   917 GAAGTTCACTCCAAGAGCCCTATACGGGCGTTCACCGCCTTGGGGACACTA 976

Qy   803 CGCTTCGGCCACGCTATGGGCGGTTGCTTCTGGCTTCCTCCAACGTGCTGCTCTC 862
         || || |||||| |||||| |    |||   ||| | |||||||||||| | |||
Db   977 CACCTCAGCCATGCATGGGGTGGCATTCTCTGTGGCCAAAGTGATGATGCTCTC 1036

Qy   863 CACGCGCCGGCCCCGCGCCTCACGGAGGCCACTGGCACCCGGACCCTTCGCGCTCTT 922
         ||||||||| |||||||||   || || || |||||||   | || |||| |||||
Db  1037 CATGCCTCGGCTGGTATGGGCTACNATGGCGCTACNCGGGCGATCTTCCAGCTGTT 1096

Qy   923 CG--GGGTCTTCGCGTCCCATCGCTTCGGCGCCGGCCTTCGGGCTAGG 979
         ||   |||| | ||      |||    || || |||||      |||||
Db  1097 GGCTCTGCTCCTTCTTCACCATCACTCCCTGTCCCGCCGTGCTGCACCTGGG 1156

Qy   980 CTCCTTCCGCGGCTCACCACTCAGTACGGCGCCGGCCGCCAACCGGCGT 1039
         ||||||| ||||||||    |  ||    ||  ||||    || || |||
Db  1157 CGCTTCTCGTGCTGCATTACCACAGGGGATCACATTTGACCACAGACT 1216

Qy  1040 CCTGTGCCTCCTTCCTGGGAGGGGGCGTGGGTGAGTCTCCAGTATGTTCGGCGAGGCTCT 1099
         ||||||||||||  |    |||    |||||||    ||||     ||||| |||||||
Db  1157 CGCTTCTCGTGCTGCATTACCACAGGGGATCACATTTGACCACAGACT 1216

Qy  1100 TCGGCACCCTTCTGGACCAAAGGCGCAAGGA 1129
         ||||  |||   |||||    ||||||||||
Db  1217 GCTGTGTGTGTCTGCTGTGGGGTTGCGGCTGCTCACAGGGCT 1276

Db  1277 GAAGGCTTTCTTCAACCAGAGTGTGGATGA 1306
```

RESULT 15
```
LOCUS       BC029819              1923 bp    mRNA    linear   PRI 30-JUN-2004
DEFINITION  Homo sapiens homolog of Drosophila Numb-interacting protein, mRNA
            (cDNA clone MGC:35310 IMAGE:5177619), complete cds.
ACCESSION   BC029819
VERSION     BC029819.1  GI:20987583
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1923)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1923)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAY-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 51 Row: O Column: 3
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 21389332.
FEATURES             Location/Qualifiers
     source          1..1923
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:35310 IMAGE:5177619"
                     /tissue_type="Brain, Lung, Testis, adult, pooled whole"
                     /clone_lib="NIH_MGC_115"
                     /note="Vector: pCMV-SPORT6"
     gene            1..1923
                     /gene="NIP"
                     /note="synonym: FLJ32334"
                     /db_xref="GeneID:90527"
     CDS             401..1852
                     /gene="NIP"
                     /codon_start=1
                     /product="Numb-interacting protein"
                     /protein_id="AAH29819.1"
                     /db_xref="GI:20987584"
```

ORIGIN

```
/db_xref="GeneID:90527"
/translation="MAILGHTPFYAGPKPTPPMDTTLASIIMIFLTALATFIVILPG
IRGKTRLPWLLRVVTSLFIGAAILAVNFSSEMSVGQVSTNTSYKAFSSEWISADIGLQ
VGLGGVNITLLTGTPVQOLNETIINYNREFTWRLGENYABEYAXALEKGLPDPVLYLAEK
FTPRSPCGLYRGYRLAGHYTSAMLWVAFLCWLLANVMLSMPVLVYGGYMLLATGIFQL
LALLFFSMATSLTSPCPLHLGASVLHTHHGPAFWIITLTTGLLCVLLGLAMAVAHRMQP
HRLKAFPNQSVDEDPMLEWSPEEGGLLSPRYRSMADSPKSQDIPLSEASSTKAYRPR
RLSIVPADVRGLAPAALSALPGALLAQAWRALLPGLRCPKAGKESRLGPPHSPWRGP
EGCEERWAEHTGDSPRPLRGRGTGRLMRWGSKERRACGVRAMLPRLVSNSGLKRPSCL
DLPKCMDYRRDARAFFHLLEPTPCVTSRHTPLI"
```

Query Match   25.1%; Score 356.4; DB 8; Length 1923;
Best Local Similarity   63.9%;
Matches 556; Conservative 0; Mismatches 311; Indels 3; Gaps 1;

```
Qy    263 CAGCCATGACCCTGTGGAACGGGCCGTACTGCCTGTTTTTACCCCAGCCCCCGGCATGCCGCAGG 322
          |||   | |||  ||  |||| || || |  ||  || ||  ||||| ||  || ||  ||
Db    397 CAAGATGGCTACTTTGGGACACATTCCCCTTCTATGCTGGCCCCAAGCCAACCTTCCC 456

Qy    323 CTTCAGCGGTTCCACTGCTCATCGTTATTCTTCAGTGTTTTTGGCTCTAGAGCAAGCTTCCT 382
          |  ||     || || |  ||  ||  | ||  |||  | |||  |   ||| |  ||
Db    397 GATGGACACCACTTTGGCCAGCATCATCATGATCTTTCTGACACTGGCCACGTTCAT 516

Qy    383 GCTCATCTTTGCCGGGGATCCGTGGCCACTTCGCGCGTGGTTTTGGTTGGGTAGAGTTCTTCT 442
          |  |   |||| | ||  || |||||| |  ||| ||||||| |||  ||  ||| || ||
Db    457 CGTCATCCTGCCTTGGCATTCGGGGGAAAAGACGAGGCTGTTCTCGGCTGCTTCGGGTGGTGAC 576

Qy    443 CAGTCTGTTCATTAGGCGCAGAAAATTGTGGCTGTGCACTTCAGTGCAGAATGGTTCGTTGGG 502
          | ||| || | |  ||  |  || |  |  |||| || || ||| ||  ||| ||| || |||
Db    517 CAGCTTATTCATCGGGGCTGCAATCCTGGCTGTGAATTTCAGTTTCGAGTGGTTCTGTGGG 636

Qy    503 TACAGTGAACACAAACACATCCTACAAAGCCTTCAGCCCAGCGCGGGTTACAGCCCGTGT 562
          | |  ||  ||  | |  ||||  |||| |||||  | ||||||  ||||  ||  |||
Db    577 CCAGGTCAGCACCAACACATCATACAAGGCCTTCAGTTCTGAGTGGGATCAGCGCTGATAT 696

Qy    563 CGGTCTGCTCGTGGGGCTGGAGGGGCCATTAATATATTCACTTCACAGGGACCCCAGTGCATCA 622
          | ||||| ||| ||||  | ||| | || ||| ||| ||| || || ||| ||||  ||  ||
Db    637 TGGGGCTGGCAGGTCGGGTGGGGTGGGAGTCAACATCACACTCACAGGGACCCCGTTGCAGCA 756

Qy    623 GCTGAACCAGACCCATTGACTACTAAACGACGCAGTTCACCTGGCGTCTGAAAGAGAATTACGC 682
          ||| |   | || ||  | ||  | ||  || ||  | || ||  ||| ||  ||||| | ||
Db    757 GCTGAAATGAGACCATCAATTACAACGAGAGGATTCACCTGGGTGAGAACTATGC 816

Qy    683 CGCGGAGTACGGCAACCGCACTCGAGAAGGGGCTGCGGCAGTGTCTTACCTGGCGCGA 742
          | || |||| ||  || | | |||||  | |  |||  ||  |  | || || | |||
Db    817 TGAGGAGTATGCAAAAGGCTCTGGAGAAGGGCTGCCAGACCCTGTTGTTGTACCTGAGCTGA 876

Qy    743 GAAGTTCACACCGAGTAGCCCTTGCGGCCCTGTACCACCAGTACCACCTGGCGGGGGACACTA 802
          |||||| ||  | || || |  ||  || ||| |||||||| ||  | |||||||||||||||
Db    877 GAAGTTCACTCCAAGAAGCCCTATACGGCCACGCGCCTTGGCGGGGACACTA 936

Qy    803 CGCCTCCGGCCACCGCTATGGCGGTTCTTGCTTCTCGGCTTCCTCTCCAACGTGCTGCTCTC 862
          ||| | ||||||| | ||||| ||| | ||  ||| |||  || ||  ||||| || || |
Db    937 CACCTCAGCCATGCTATGGGTGGCATTCCTCTCGCTGGCCAATGTGATGCTCTC 996

Qy    863 CACGCCGCGCCCCCCTCTCACGGAGGCGCACTGCTGACCACACGGAGCCTTGCGCTCTT 922
          ||   |||| ||||||| |||| ||| | | |  ||| || ||  || || ||  || |
Db    997 CATGCCCTGTGCTGGTATATAGGTGGCTACATGCTACTGGCCACCGGGATCTTCCAGCTGTT 1056

Qy    923 CG---GGGTCTTCGCCTTGGCCTTCCATCTCAGGGTGCGGCGTCTTGCCGCCTAGG 979
          ||   ||| | || |  ||||  || |||| |  || ||  || | ||||  ||||
Db   1057 GGCTCTGCTCTTCTTTTCCATGGCCACATCACTCACCTCTCCCTGCCCGCCCCTTGG 1116

Qy    980 CTCCTCCGGCTCACCACTCAGTACGGCGCGCCGCCGCCTTCTGGGTGCGTGGCAACCGGGGGT 1039
          || ||  ||| |||| | ||| |||||| |  | ||  ||||| |||| |||  |  ||||||
Db   1117 CGCTTCTGTGCGCATACTCACCATGGGCCCTGCCTTTCTGGGATCACCATTGACCACAGAGACT 1176

Qy   1040 CCTGTGCCTCTTCCTCGGAGGGCCGTGGTGAGTCTCCAGTATGTTCGGCCCCAGCCGCTCT 1099
          || || | || ||||||||||  |||| |||| |  |||| |||  ||||||| || || |
Db   1177 GCTGTGTGCTGCTGGGCGTGGGGCCTATGCGGTGGCCCACAGAGATGCAGCCCTCACAGAGCT 1236

Qy   1100 TCGCACCCTTCTGGACCAAAGCGCCAAGGA 1129
          | | |||||| | |||||||| ||||||||
Db   1237 GAAGGCTTTCTTCAAACCAGAGTGTGGGATGA 1266
```

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 19:32:01

Copyright (c) 1993 - 2006 Compugen Ltd.

GenCore version 5.1.6

; Search time 937 Seconds
(without alignments)
10107.287 Million cell updates/sec

Title: US-09-937-059-57
Perfect score: 1421
Sequence: 1 gaaccaggaaagtaacggct.........ctttttttttcttgttttt 1421

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	ID	Description
1	1421	100.0	1421	3	AAA96505	Aaa96505 cDNA enco
2	1421	100.0	1450	12	ADL06425	Adl06425 Human tum
3	1418.6	99.8	1474	6	ABL90535	Abl90535 Human pol
4	1413	99.4	1441	3	AAA94623	Aaa94623 Human CAS
5	1412	99.4	1491	10	ADD19228	Add19228 Human cDN
6	1375.8	96.8	1460	8	AAC95564	Aac95564 Human sec
7	1375.8	96.8	1460	10	ABZ73640	Abz73640 Secreted
8	1375.8	96.8	1460	8	ADC20289	Adc20289 Human sec
9	1375.8	96.8	1460	10	ABZ67235	Abz67235 Human sec
10	958.4	67.4	960	13	ADU01671	Adu01671 Novel hum
11	768.8	54.1	981	6	ACF58262	Acf58262 Human chr
12	768.8	54.1	981	12	ADO85846	Ado85846 Colon can
13	453	31.9	498	3	AAA96624	Aaa96624 Human CAS
14	413.4	29.1	5033	8	ABZ74587	Abz74587 Secreted
15	413.4	29.1	5033	10	ADC21005	Adc21005 Human sec
16	413.4	29.1	5033	10	ABZ68115	Abz68115 Human sec
17	413	29.1	580	10	ADD19187	Add19187 Human cDN
18	402.8	28.3	406	5	AAS66563	Aas66563 DNA encod
19	391	27.5	2684	10	ACF58261	Acf58261 ENST00000
c 20	391	27.5	2684	10	ADA53627	Ada53627 Human cod
c 21	391	27.5	2684	12	ADN06056	Adn06056 Antipsori
c 22	391	27.5	2684	12	ADO85844	Ado85844 Colon can
23	356.4	25.1	1636	6	ABN59923	Abn59923 Novel hum
24	354.8	25.0	1029	5	AAF45134	Aaf45134 Human INT
25	354.8	25.0	1029	8	ACD66787	Acd66787 Secreted
26	354.8	25.0	1029	12	ADQ10402	Adq10402 Human pol
27	354.8	25.0	1029	12	ADQ98274	Adq98274 Human pol
28	354.8	25.0	1029	5	AAF45133	Aaf45133 Human INT
29	354.8	25.0	1029	12	ABX94107	Abx94107 cDNA enco
30	354.8	25.0	1029	8	ACD66786	Acd66786 Secreted
31	354.8	25.0	2133	12	ADQ10401	Adq10401 Human pol
32	354.8	25.0	2133	12	ADQ98273	Adq98273 Human pol
33	348	24.5	348	10	ADG32833	Adg32833 Human DNA
34	348	24.5	348	12	ADO85843	Ado85843 Colon can
35	299.4	21.1	2439	10	ACF58261	Acf58261 Human cod
36	245.8	17.3	2684	10	AAS67707	Aas67707 DNA encod
37	245.8	17.3	2684	12	ADO85844	Ado85844 Colon can
38	245.8	17.3	2684	12	ADA53627	Ada53627 Human cod
39	161.8	11.4	1338	5	ABL07775	Abl07775 Drosophil
40	161.1	11.4	1602	4	ABL21213	Abl21213 Drosophil
41	65.6	4.6	512	4	AAH36185	Aah36185 Human col
42	65.6	4.6	512	4	AAH36185	Aah36185 Human col
c 43	64	4.5	512	4	ABL07775	Abl07775 Drosophil
c 44	60.6	4.3	2000	8	ADA71938	Ada71938 Rice gene
c 45	52.6	3.7	134	8	ABZ74588	Abz74588 Secreted

ALIGNMENTS

RESULT 1
AAA96505

ID   AAA96505 standard; cDNA; 1421 BP.
XX
AC   AAA96505;
XX
DT   08-FEB-2001 (first entry)
XX
DE   cDNA encoding a human transmembrane protein.
XX
KW   Human; transmembrane protein; cell proliferation disorder; myeloma;
KW   reproductive disorder; smooth muscle disorder; neurological disorder;
KW   arteriosclerosis; leukaemia; acquired immunodeficiency syndrome; AIDS;
KW   allergy; ovulatory defect; angina; hypertension; stroke; epilepsy;
KW   Alzheimer's disease; Tourette's disorder; SS.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   267..1229
FT   /*tag= a
XX
PN   WO200056891-A2.
XX
PD   28-SEP-2000.
XX
PF   22-MAR-2000; 2000WO-US007817.
XX
PR   22-MAR-1999; 99US-0125537P.
PR   16-JUN-1999; 99US-0139565P.
XX
DR   WPI; 2000-579485/54.
DR   P-PSDE; AAB18992.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Yue H, Lal P, Tang YT, Hillman JL, Reddy R, Bandman O;
PI   Baughn MR, Lu DAM, Azimzai Y, Yang J;
XX
PT   New human transmembrane proteins are used to treat a disease or condition
PT   associated with decreased expression of functional HTMP e.g. Tourette's
PT   disorder, angina and leukemia.

```
XX
PS   Claim 4; Page 129; 130pp; English.
XX
CC   The present sequence encodes a human transmembrane proteins (HTMP).
CC   Agonists and antagonists of the protein are used to treat a disease or
CC   condition associated with overexpression of the protein. Diseases and
CC   conditions which can be treated include cell proliferative,
CC   immunological, reproductive, smooth muscle and neurological disorders
CC   e.g. arteriosclerosis, myeloma, leukaemia, acquired immunodeficiency
CC   syndrome (AIDS), allergies, ovulatory defects, angina, hypertension,
CC   stroke, Alzheimer's disease, epilepsy and Tourette's disorder. The
CC   polynucleotides may be used to detect and quantify gene expression in
CC   biopsied tissues where protein expression may be correlated with disease
CC   e.g. to determine absence, presence or excess expression of HTMP or to
CC   monitor regulation of HTMP expression during therapeutic intervention
XX
SQ   Sequence 1421 BP; 264 A; 468 C; 395 G; 294 T; 0 U; 0 Other;

Query Match            100.0%;  Score 1421;  DB 3;  Length 1421;
Best Local Similarity  100.0%;  Pred. No. 0;  Mismatches 0;  Indels 0;  Gaps 0;
Matches 1421;  Conservative 0;
```

```
Qy      1 GAACCAGGAAAAGTAAACGGCTACAGACACTGAGAAATAGTTTCCGCTCGCGCGGCTAGAAAAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAACCAGGAAAAGTAAACGGCTACAGACACTGAGAAATAGTTTCCGCTCGCGCGGCTAGAAAAA 60

Qy     61 CTCTGTCTCGGTACCAACCCCAGAGACGGTTGAGAGCAGAGCCACCTCCAGCGTTCCTTAACGGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCTGTCTCGGTACCAACCCCAGAGACGGTTGAGAGCAGAGCCACCTCCAGCGTTCCTTAACGGA 120

Qy    121 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCCTTGTACGCAAAGAGAACG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCCTTGTACGCAAAGAGAACG 180

Qy    181 CCAAAGGACGACGCTCTCCCGGCGTTCCAGGCCAGCCCACTCGGTCCCAGCCTTGTACGCAAAGAGGACG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCAAAGGACGACGCTCTCCCGGCGTTCCAGGCCAGCCCACTCGGTCCCAGCCTTGTACGCAAAGAGGACG 240

Qy    241 CGTGCAGCACTCGGCCGCGGCGTGAAGCATGACCCGTGTGGAAACGGCGTACTGCCTTTTTTTACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGTGCAGCACTCGGCCGCGGCGTGAAGCATGACCCGTGTGGAAACGGCGTACTGCCTTTTTTTACC 300

Qy    301 CCCAGCCCCGGCCATGCCCGGCGTTCAGCGGTTCCATCGTGTATTCTTAGTGTGTTTT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCCAGCCCCGGCCATGCCCGGCGTTCAGCGGTTCCATCGTGTATTCTTAGTGTGTTTT 360

Qy    361 TGGCTCTCTAGCAGCGAAGCTTCCTTGCTCTATCTTTGCGGGATCCGTGGCGCACTCGCGCTGGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGGCTCTCTAGCAGCGAAGCTTCCTTGCTCTATCTTTGCGGGATCCGTGGCGCACTCGCGCTGGT 420

Qy    421 TTTGGTTGGTGAGAGCTTCTTCTTCAGTCTCGTTCATAGGGCGCAGAAATTGTGGCGTGTGCACT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTTGGTTGGTGAGAGCTTCTTCTTCAGTCTCGTTCATAGGGCGCAGAAATTGTGGCGTGTGCACT 480

Qy    481 TCAGTGCAGAATGGTTCGTTGGGGTACAGTGAAACCACAACATTCTACAAAAGCCTTCAGCG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCAGTGCAGAATGGTTCGTTGGGGTACAGTGAAACCACAACATTCTACAAAAGCCTTCAGCG 540

Qy    541 CAGGCGCCGGTTACAGCCCGTGTCGGGTCTGTCTGCCTGGAGGGCCATTAATATTACAC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CAGGCGCCGGTTACAGCCCGTGTCGGGTCTGTCTGCCTGGAGGGCCATTAATATTACAC 600

Qy    601 TCACAGGGAACCCCAGTGCATCAGCTGAACGTGACCAATTGACTACAACGAGCAGTTCACTT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCACAGGGAACCCCAGTGCATCAGCTGAACGTGACCAATTGACTACAACGAGCAGTTCACT 660

Qy    661 GGGGTCTTGAAAAGACGAATTACGCCCCGGAGTACGCCCGGAGTACGCCCGGAGTACGCCCGG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGGGTCTTGAAAAGACGAATTACGCCCCGGAGTACGCCCGGAGTACGCCCGGAGTACGCCCGG 720

Qy    721 ACCCAGTGCTCTACCTGGCGGGAGAAGTTCACACCGAGTAGCCCTTGCGCGCCTGTACCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACCCAGTGCTCTACCTGGCGGGAGAAGTTCACACCGAGTAGCCCTTGCGCGCCTGTACCACC 780

Qy    781 AGTACCACCTGGGCGGGACAACTACCGCGCTCGGCCGCCACGGCTATGGGGTGGGGTGCGGTTCGCTTGCTTGGC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGTACCACCTGGGCGGGACAACTACCGCGCTCGGCCGCCACGGCTATGGGGTGGGGTGCGGTTCGCTTGCTTGGC 840

Qy    841 TCCCTCTCAAACGTGCTGCTTCCACGCGCCGGCCCCGCTTCTACGGAGGCCGGACCACTGCTGA 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCCCTCTCAAACGTGCTGCTTCCACGCGCCGGCCCCGCTTCTACGGAGGCCGGACCACTGCTGA 900

Qy    901 CCACCGGAGCCGTTCGCGCTCTTCGGGGGTCTTCGGCCTTGCCCTTCCATTCTCTTAGCCTCCGCG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCACCGGAGCCGTTCGCGCTCTTCGGGGGTCTTCGGCCTTGCCCTTCCATTCTCTTAGCCTCCGCG 960

Qy    961 TCTGCCCGCTCCGCGCCTAGCCTCTCCGCGCGCTCCTCCGCGCTCACCCACTCAGTAACGGTACGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TCTGCCCGCTCCGCGCCTAGCCTCTCCGCGCGCTCCTCCGCGCTCACCCACTCAGTAACGGTACGGGG 1020

Qy   1021 TCACGCCTGGCAACCCGGCGTCCTGGTGCCCTCCTTCCTCTGGAGGGGCCGTGGTGAGTCTCCAGT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCACGCCTGGCAACCCGGCGTCCTGGTGCCCTCCTTCCTCTGGAGGGGCCGTGGTGAGTCTCCAGT 1080

Qy   1081 ATGTTCGGCGCCCAGGCGCTCTTCGGCACCCCTTCTCGGACCAAAGCGCCAAAGGACTGCAGCCAGG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ATGTTCGGCGCCCAGGCGCTCTTCGGCACCCCTTCTCGGACCAAAGCGCCAAAGGACTGCAGCCAGG 1140

Qy   1141 AGAGAGGGGGGGCTCACCTCTTATCCTCGGGGGAGGGGGGGGACCCCACTGACGCAACGCTCTCCCAG 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AGAGAGGGGGGGCTCACCTCTTATCCTCGGGGGAGGGGGGGGACCCCACTGACGCAACGCTCTCCCAG 1200

Qy   1201 ACTTAAAATGTATCACCACTAACCTGTGAGGGGGACCCAAATCTGGACTTCCTTCCCGGCCT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACTTAAAATGTATCACCACTAACCTGTGAGGGGGACCCAAATCTGGACTTCCTTCCCGGCCT 1260

Qy   1261 TGGGACAATCGGCCAGGCCGGGAAGCAGTGCCCGGCCAGGGCTCCAGGAACTCCAGGAA 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGGGACAATCGGCCAGGCCGGGAAGCAGTGCCCGGCCAGGGCTCCAGGAACTCCAGGAA 1320

Qy   1321 GGGCACTGAGCGCTGCTGGCGCGGGGAGCGACATCCGGACCAGGAAAAGTCTCC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GGGCACTGAGCGCTGCTGGCGCGGGGAGCGACATCCGGACCAGGAAAAGTCTCC 1380

Qy   1381 TGGGGCGGATCTGTAAATAAACCTTTTTTTTTTTCTTTTTGTTTTTTT 1421
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGGGCGGATCTGTAAATAAACCTTTTTTTTTTTCTTTTTGTTTTTTT 1421
```

```
RESULT 2
ADL06425
ID   ADL06425 standard; cDNA; 1450 BP.
XX
AC   ADL06425;
XX
DT   20-MAY-2004 (first entry)
XX
DE   Human tumour-associated antigenic target (TAT) cDNA sequence #5 .
XX
KW   Human; tumour-associated antigenic target; TAT; cell death; tumour;
KW   cancer; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004016225-A2.
XX
PD   26-FEB-2004.
XX
PF   19-AUG-2003; 2003WO-US025892.
XX
PR   19-AUG-2002; 2002US-0404809P.
PR   21-AUG-2002; 2002US-0405645P.
PR   23-SEP-2002; 2002US-0413192P.
PR   15-OCT-2002; 2002US-0419008P.
PR   15-NOV-2002; 2002US-0426847P.
PR   02-JUL-2003; 2003US-0484959P.
```

```
XX
PA   (GBTH ) GENENTECH INC.
XX
PI   Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;
PI   Spencer SD, Wu TD, Zhang Z;
XX
DR   WPI; 2004-257144/24.
DR   P-PSDB; ADL06505.
XX
PT   New antibody that binds to a tumor-associated antigenic target (TAT)
PT   polypeptide, useful for preparing a composition for diagnosing or
PT   treating cancer.
XX
PS   Claim 1; SEQ ID NO 5; 319pp; English.
XX
CC   The present invention relates to the isolation of human tumour-associated
CC   antigenic target (TAT) polynucleotide and polypeptide sequences. Also
CC   disclosed is an antibody that binds to a TAT polypeptide. The antibody is
CC   a monoclonal antibody, an antibody fragment, a chimeric antibody or a
CC   humanised antibody. It is conjugated to a growth inhibitory agent, a
CC   produced in bacteria or in CHO cells and induces death of a cell to which
CC   it binds. The antibody is useful for preparing a composition for
CC   diagnosing or treating tumours and cancer. The present sequence
CC   represents a human TAT cDNA sequence of the invention.
XX
SQ   Sequence 1450 BP; 475 A; 279 C; 398 G; 297 T; 0 U; 1 Other;

Query Match             100.0%;  Score 1421;  DB 12;  Length 1450;
Best Local Similarity   100.0%;  Pred. No. 0;  Mismatches 0;
Matches 1421;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 GAACCAGGAGAAAGTAACGGCTACAGACAGTGAGAAAATAGTTCGTCGCCGGCTAGAAAAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9 GAACCAGGAGAAAGTAACGGCTACAGACAGTGAGAAAATAGTTCGTCGCCGGCTAGAAAAA 68

Qy     61 CTCTGTCGGTACCAACCCCAGAGAGCGGTTGAGAGCAGCACCTCCAGCGCTTCCTTAACGGA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 CTCTGTCGGTACCAACCCCAGAGAGCGGTTGAGAGCAGCACCTCCAGCGCTTCCTTAACGGA 128

Qy    121 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCACTCGGTTGTAGGCAAAGAGACG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCACTCGGTTGTAGGCAAAGAGACG 188

Qy    181 CCAAGGACGCGCTCTCCCGCGTCAGGCAGCCCAGCGTTGCTGGCTTGCCTGCCCGCCTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 CCAAGGACGCGCTCTCCCGCGTCAGGCAGCCCAGCGTTGCTGGCTTGCCTGCCCGCCTG 248

Qy    241 CGTGCAGCACTCCGCCGCGGGGTCGAGCATGACCCGTGGAACGGCGCTACTGCCTTTTTACC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    249 CGTGCAGCACTCCGCCGCGGGGTCGAGCATGACCCGTGGAACGGCGCTACTGCCTTTTTACC 308

Qy    301 CCCAGCCCCCGGCCATCGCCGCAGGCGTTCCACTGCTCATCGTTATTCTCTAGTGTTTTT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    309 CCCAGCCCCCGGCCATCGCCGCAGGCGTTCCACTGCTCATCGTTATTCTCTAGTGTTTTT 368

Qy    361 TGGCTCTAGCAGCAAGCTTCCTGCTCTCATCTCATCTTGCCGGGGATCCGGCCCACTCGGGCTGGT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    369 TGGCTCTAGCAGCAAGCTTCCTGCTCTCATCTCATCTTGCCGGGGATCCGGCCCACTCGGGCTGGT 428

Qy    421 TTTGGTTGGTTGAGAGCTTCTTCTCAGTCTGTTCATAGGCGGAGAAATTGTGTGGCTGCACT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    429 TTTGGTTGGTTGAGAGCTTCTTCTCAGTCTGTTCATAGGCGGAGAAATTGTGTGGCTGCACT 488

Qy    481 TCAGTGCAGAAATGGTTCGTGGGGTAACACAACACATCCTACAGAAAGCCTTCAGCG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    489 TCAGTGCAGAAATGGTTCGTGGGGTAACACAACACATCCTACAGAAAGCCTTCAGCG 548

Qy    541 CAGCGCCGGTTACAGCCCGTGTCGGGTCGTCGTGGGTCCTGGAGGGCCATTAATATTACAC 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    549 CAGCGCCGGTTACAGCCCGTGTCGGGTCGTCGTGGGTCCTGGAGGGCCATTAATATTACAC 608

Qy    601 TCACAGGGACCCCAGTGCATCAGCTGAACCAGGACCAATTGACTACAAACGACCAGTTCACCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    609 TCACAGGGACCCCAGTGCATCAGCTGAACCAGGACCAATTGACTACAAACGACCAGTTCACCT 668

Qy    661 GGCGTCTGAAAGAGATAATTACGCCCGGGGAGTAGCGGAACGGCACTGGAGAAGGGGCCG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    669 GGCGTCTGAAAGAGATAATTACGCCCGGGGAGTAGCGGAACGGCACTGGAGAAGGGGCCG 728

Qy    721 ACCCAGTGCTCTACCTTGCCGGAGGAAGGTTCACACCGAGTACCCAGTTGCCGGCCGTGACCACC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    729 ACCCAGTGCTCTACCTTGCCGGAGGAAGGTTCACACCGAGTACCCAGTTGCCGGCCGTGACCACC 788

Qy    781 AGTACCCACCTGGGCCGGGGACTACCGCCTTCGCGGCTATGGGGTGGCCGGTTCGTTCTTCTGGC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    789 AGTACCCACCTGGGCCGGGGACTACCGCCTTCGCGGCTATGGGGTGGCCGGTTCGTTCTTCTGGC 848

Qy    841 TCCTCTCCAACGTGCTGCTTCCTCCAGCCCGGCGCCCGCTCTACGGAGGCCGGCACTGCTGA 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    849 TCCTCTCCAACGTGCTGCTTCCTCCAGCCCGGCGCCCGCTCTACGGAGGCCGGCACTGCTGA 908

Qy    901 CCACCGGAGCCGCCTTCGCGCTCTTCGGGGTGCTTCTTGGCCTTCCATCTCTAGCGTCGCGC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    909 CCACCGGAGCCGCCTTCGCGCTCTTCGGGGTGCTTCTTGGCCTTCCATCTCTAGCGTCGCGC 968

Qy    961 TCTGCCCGCGTCCGGCCTAGGGTCTCCGCGTCACCACTCACTCAGTACGGCGCGCCCTTCTGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    969 TCTGCCCGCGTCCGGCCTAGGGTCTCCGCGTCACCACTCACTCAGTACGGCGCGCCCTTCTGGG 1028

Qy   1021 TCACGCTGGCAACCGGCGTGCCTGTGGCCTCTTCCTCGGAGGGGCCGGGTGAGTCTCCAGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1029 TCACGCTGGCAACCGGCGTGCCTGTGGCCTCTTCCTCGGAGGGGCCGGGTGAGTCTCCAGT 1088

Qy   1081 ATGTTCGGCCCCAGCGGCTCTTCGCACCCCTTCTGGACCAAAGCGCCAAGGACTGCAGCGAGG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1089 ATGTTCGGCCCCAGCGGCTCTTCGCACCCCTTCTGGACCAAAGCGCCAAGGACTGCAGCGAGG 1148

Qy   1141 AGAGAGGGGGGCTCACCTCTTATCCTCGGCGAACAGCAGCAGCAGCCGCTCTCTCCAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1149 AGAGAGGGGGGCTCACCTCTTATCCTCGGCGAACAGCAGCAGCAGCCGCTCTCTCCAG 1208

Qy   1201 ACTTAAAATGTATCACCACTAACCTGTGAGGGGGAGACCAATCTGGACTCCTTCCCGCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1209 ACTTAAAATGTATCACCACTAACCTGTGAGGGGGAGACCAATCTGGACTCCTTCCCGCCT 1268

Qy   1261 TGGGACATCCGCAGGCGCGGGGAAGCACGTGCCCGGCCGCCCAGGAGAGACTCCAGGAA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1269 TGGGACATCCGCAGGCGCGGGGAAGCACGTGCCCGGCCGCCCAGGAGAGACTCCAGGAA 1328

Qy   1321 GGGCACTGAGGGCGTGCTGGGCCGGAGGGCCTCGGACATCCGGCAGGCACCACGAAAGTCTCC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1329 GGGCACTGAGGGCGTGCTGGGCCGGAGGGCCTCGGACATCCGGCAGGCACCACGAAAGTCTCC 1388

Qy   1381 TGGGGCGGATCTGTAAATAAACCTTTTTTTTTCTTTTGTTTTT 1421
          |||||||||||||||||||||||||||||||||||||||||||
Db   1389 TGGGGCGGATCTGTAAATAAACCTTTTTTTTTCTTTTGTTTTT 1429
```

```
PN   WO20190304-A2.
XX
PD   29-NOV-2001.
XX
PF   18-MAY-2001; 2001WO-US016450.
PR   19-MAY-2000; 2000US-0205515P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE, Rosen CA;
XX
DR   WPI; 2002-122018/16.
DR   P-PSDB; ABB90126.
XX
PT   Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT   prevention of neural, immune system, muscular, reproductive,
PT   gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT   disorders.
XX
PS   Claim 4; SEQ ID NO 1097; 2081pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC   (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC   breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ;(e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1474 BP; 292 A; 480 C; 404 G; 296 T; 0 U; 2 Other;

Query Match             99.8%;  Score 1418.6;  DB 6;  Length 1474;
Best Local Similarity   99.8%;  Pred. No.0:  1;  Indels   0;  Gaps   0;
Matches 1418;  Conservative   2;  Mismatches 0;
```

```
Qy         1 GAACCAGGAAAGTAAACGGTACAGACAGTGAGAAATAGTTTTCGCTCGCGGCCTAGAAAAA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TTTGGTTGGTTGGAGAGTTCTTCTCAGTTCTGTTTCATTAGGGGCCAGAAATTTGGGCTGCACT 480

Qy        19 GAACCAGGAAAGTAAACGGTACAGACAGTGAGAAATAGTTTTCGCTCGCGGCCTAGAAAAA 78
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       439 TTTGGTTGGTTGGAGAGTTCTTCTCAGTTCTGTTTCATTAGGGGCCAGAAATTTGGGCTGCACT 498

Qy        61 CTCTGTCGGTACCAACCCCAGAGGAGGCGGTTGAGAGCAGGCCACCTCCACCTTCCTTAACGGA 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TCAGTGCAGAATGGTTCGTTGGGGTACAGTGAACACCACCAACACCATCCTTACACAGCCTTCAGCG 540

Qy        79 CTCTGTCGGTACCAACCCCAGAGCGTTGAGAGCAGCCCACCTCCACGCGTTCCTTAACCGA 138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       499 TCAGTGCAGAATGGTTCGTTGGGGTACAGTGAACACCAACACCATCCTTACAGCCTTCAGCG 558

Qy       121 GAGGTGCAGGACTTCAGGACTTCACCAGGCCCACTCGGGTCCCAGCCAAAGAGAGACG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 CACGCGCGGGCGTTACAGCCCGTGTCGTCGGTCTTGGGGCCTGTGGGCCTTGGGAGGGGCACG 600

Qy       139 GAGGTGCAGGACTTCAGACTTCACCAGGCCCACTCGGGTCCCAGCCAAAGAGAGACG 198
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db       559 CAGGCGCGGCGGTTACAGCCCGTGTCCGGTCTCGGGTCTCGGGGCCTGGGCTTCACACT 618

Qy       181 CCAAGGACGCGCTCTCCCGCGTTCCAGGCAGCCCCAGCCTTGCTGGCTTGCTGGCCCGCCTG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 TCACAGGGACCCCAGTGCATCAGCTGACGTGACGAGACGAGACAGACTTGACTCACACAGACAGTTCACCT 660

Qy       199 TCAAGGACGCGCTCTCCCGCGTTCCAGGCAGCGCCCAGCCTTGCTTGCTTGCCCGCCTG 258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       619 TCACAGGGACCCCAGTGCATCAGCTGACGTGACGAGACGAGACCAGACCAGCCTTGCTCGCGG 678

Qy       241 CGTGCAGCACTCGGCCGCGGCCGTGCAGACATGACCGGTACTACACCGGTTACC 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       679 GGGCGTCTGAAAGAGAGAAGAATTACGCCCGCGACTACGCCGGAACGGCACTGGAGGAAGGGCTGCCG 738

Qy       259 CGTGCAGCACTCCGCCGCGGCCGTGCAGACATGACCGGTACTACACCGGTTACC 318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       699 AGTACCACCTGTCGGGGGGGACACTACGGGCTCGGCCGCCTCGGCCACGCTAGGGTGGCCTTCAGCC 798

Qy       301 CCCAGCCCCGGACCCCGGGCCATGCCCAGGGCTTCAGCGGTTCAGCGGTTTT 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       739 ACCCAGTGCTTCTACCCTGGGGGGAGAAGTTCACACCGGAGTAGCCCTTGCGGCCGTGTACCACC 780

Qy       319 CCCAGCCCCGGGCCATGCCCGGCAGGGCTTCAGCGGTTATTCTAGTGTTTT 378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       759 AGTACCACCTGTCGGGGGGGACACTACGGCTTGGGCCGCCTCGGCCGCCTATGGGTGGCGTTCGTGGC 840

Qy       361 TCGGCTCTAGCAGCAAGCTTCCTGCTCATCTTCATGCGGGGATCCGGTGT 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       799 AGTACCACCTGGGCGGGGACACTACGGCCTCGGCCGCCTCGGCCACGCTATGGGTGGCTTCGTCGGGC 858

Qy       379 TGGCTCTCAGCAGCAAGCTTCATCTTTCRCGGGGATCATCGTTATTCTAGTGTTT 438
             ||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       859 TCCTCTCCAAGCGTGCTGCTCTCCACGCCCGGCCCCCGCTTCTACGGAGGCTTCCATCTCTAGGGTTGCGC 918
```

```
RESULT 4
AAA94623
ID   AAA94623 standard; DNA; 1441 BP.
```

```
XX
AC   AAA94623;
XX
DT   11-JAN-2001  (first entry)
XX
DE   Human CASB618 coding sequence.
XX
KW   Epitope; human; CASB618; chromosome 15; tumour; vaccine; cancer; ovarian;
KW   colon; autoimmune disease; HLA_A0201; 88.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             259..1221
FT                   /*tag= a
FT                   /product= "Human CASB618"
XX
PN   WO200053748-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002048.
XX
PR   11-MAR-1999;  99GB-00005607.
PR   01-SEP-1999;  99GB-00020590.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Bruck CEM, Cassart J, Coche T, Vinals Y De BassoleC;
XX
DR   WPI; 2000-572268/53.
DR   P-PSDB; AAB26325.
XX
PT   New human CASB618 polypeptide, useful as a vaccine for prophylactic and
PT   therapeutic treatment of cancers, particularly ovarian or colon cancer,
PT   autoimmune diseases and related conditions.
XX
PS   Claim 13; Page 61; 76pp; English.
XX
CC   The present sequence is the coding sequence of human CASB618 protein. The
CC   gene for human CASB618 is thought to be located on chromosome 15. The
CC   protein encoded by the present sequence and epitopes of the CASB618
CC   protein (see AAB26327 to AAB26399) are useful in diagnosing the occurence
CC   of tumour cells and in vaccines for prophylactic and therapeutic
CC   treatment of cancers, particularly ovarian or colon cancer, autoimmune
CC   diseases and related conditions
XX
SQ   Sequence 1441 BP; 289 A; 466 C; 392 G; 294 T; 0 U; 0 Other;
```

```
Query Match             99.4%;  Score 1413;  DB 3;  Length 1441;
Best Local Similarity   100.0%;  Pred. No. 0;  Matches 1441;
Matches 1413;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy          9 AAAGTAACGGCTACAGACAGTGAGAAATAGTTTCGCTCCGCCGCTAGAAAAAACTCTGTCG 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAGTAACGGCTACAGACAGTGAGAAATAGTTTCGCTCCGCCGCTAGAAAAAACTCTGTCG 60

Qy         69 GTACCAACCCCAGAGCAGGGTTGAGAGAGGCAGCCACCCACTCCACGGTTCCTTTTCGGTTG 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTACCAACCCCAGAGCAGGGTTGAGAGAGGCAGCCACCCACTCCACGGTTCCTTTTCGGTTG 120

Qy        129 GGACTCAGAGCTTCACCAGCCCACTCGGTCCCAGAGGGGGGCAAAGAGACGGCCAAGGAC 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGACTCAGAGCTTCACCAGCCCACTCGGTCCCAGAGGGGGGCAAAGAGACGGCCAAGGAC 180

Qy        189 GCGCTCTCCGGCGTTCCAGGCAGCGCCCACTCGGTTCGGCTTGCGGGCGTTGCGGGCAGC 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCGCTCTCCGGCGTTCCAGGCAGCGCCCACTCGGTTCGGCTTGCGGGCGTTGCGGGCAGC 240

Qy        249 ACTCGGGCCGGCGGCGTGCAGCATGAACCCTGTGAACGGCGTACTGCCTTTTTTACCCCCAGCCC 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTCGGGCCGGCGGCGTGCAGCATGAACCCTGTGAACGGCGTACTGCCTTTTTTACCCCCAGCCC 300

Qy        309 CGGCCATGCCGCCCAGGGGCTTCAGCGTTCCACTGCTCGTTATTCTAGTGTGTTTGGCTCTA 368
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGGCCATGCCGCCCAGGGGCTTCAGCGTTCCACTGCTCGTTATTCTAGTGTGTTTGGCTCTA 360

Qy        369 GCAGCAAGCTTCCTTGCTCATCTTGCCGGGGGATCCCGTAATCGTTGATTCTGGCTGGTTG 428
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGCAAGCTTCCTTGCTCATCTTGCCGGGGGATCCCGTAATCGTTGATTCTGGCTGGTTG 420

Qy        429 GTGAGGAGTTCTTCTCAGTCTGTTCATAGGCCGCAGAAATTGTGGCTGTGCACTTCAGTGCA 488
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGAGGAGTTCTTCTCAGTCTGTTCATAGGCCGCAGAAATTGTGGCTGTGCACTTCAGTGCA 480

Qy        489 GAATGGTTTCGTGGGGTACAGTGAACACCAACACAACATCCTACAGAGCCTTCAGCAGCGGC 548
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAATGGTTTCGTGGGGTACAGTGAACACCAACACAACATCCTACAGAGCCTTCAGCAGCGGC 540

Qy        549 GTTACAGCCCGTGGTCGGGTCGGGTCTTGCTCGTCGTTGGGGGCAGTTAAATTACACTTCACAGGG 608
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTTACAGCCCGTGGTCGGGTCGGGTCTTGCTCGTCGTTGGGGGCAGTTAAATTACACTTCACAGGG 600

Qy        609 ACCCCAGTGCATCAGCTGAACGAGCACCATTGACTACACAACGAGCAGTTCACCTGGCGTCTG 668
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCCCAGTGCATCAGCTGAACGAGCACCATTGACTACACAACGAGCAGTTCACCTGGCGTCTG 660

Qy        669 AAAGAGAAATTACGCGCGCGGAGTACCGCGAAGGCACTGGGAGAAAGGGCGGGCCACCCCAGTG 728
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAAGAGAAATTACGCGCGCGGAGTACCGCGAAGGCACTGGGAGAAAGGGCGGGCCACCCCAGTG 720

Qy        729 CTCTACCTGGCGGGGAGGAAGTTCACACGGAGTAGCCCTTGCGGCCGTGTACCACCAGTACCAC 788
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTACCTGGCGGGGAGGAAGTTCACACGGAGTAGCCCTTGCGGCCGTGTACCACCAGTACCAC 780

Qy        789 CTGGCCGGGGACAACTACCGCCTCGGCCACGCTATGGGTGGGTGGTCGCGTTCGCTTCGGCTCTCTCC 848
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGGCCGGGGACAACTACCGCCTCGGCCACGCTATGGGTGGGTGGTCGCGTTCGCTTCGGCTCTCTCC 840

Qy        849 AACGTTCGTGCTCTCCACGCGCCCCGCGGCTTCACGGAGGCCTGGCCACTGCTGACCACCGGA 908
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACGTTCGTGCTCTCCACGCGCCCCGCGGCTTCACGGAGGCCTGGCCACTGCTGACCACCGGA 900

Qy        909 GCCTTCGGCGTCTTCGGGGGGTTCCGGCCTTGGCTCCCATCTCTAGCGTGGCCCCTCGCCCG 968
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCCTTCGGCGTCTTCGGGGGGTTCCGGCCTTGGCTCCCATCTCTAGCGTGGCCCCTCGCCCG 960

Qy        969 CTCCGGCTAGGCTCCTCCGCGCCTCACCACTCAGTACGGGCCCTTCTGGGGTCACGGCTG 1028
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTCCGGCTAGGCTCCTCCGCGCCTCACCACTCAGTACGGGCCCTTCTGGGGTCACGGCTG 1020

Qy       1029 GCAACCGGCGGTCCCTGTGGCCTCTTCCTCTGGAGGGGGGCCGGTGTGAGTCTCCAGTAGTTCGG 1088
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCAACCGGCGGTCCCTGTGGCCTCTTCCTCTGGAGGGGGGCCGGTGTGAGTCTCCAGTAGTTCGG 1080

Qy       1089 CCCAGGGCTCTTCGGCCACCCTTCTGGACCAAAGGCCAAGGACTGCAGCAGGAGAGAGGG 1148
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCCAGGGCTCTTCGGCCACCCTTCTGGACCAAAGGCCAAGGACTGCAGCAGGAGAGAGGG 1140

Qy       1149 GGCTCACCTCTTATCCTCGGCGACCCACTGCAACGGCAGGCGGCGCTCTCCCAGACTTAAA 1208
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGCTCACCTCTTATCCTCGGCGACCCACTGCAACGGCAGGCGGCGCTCTCCCAGACTTAAA 1200

Qy       1209 TGTATCACCCACTAAACCTGTAGAGGGGGACGACCCAAATCTGGACTCCTTCCCCGCCTTGGGACAT 1268
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGTATCACCCACTAAACCTGTAGAGGGGGACGACCCAAATCTGGACTCCTTCCCCGCCTTGGGACAT 1260

Qy       1269 GCGAGGCCGGGAAGCAGTGCCCGCGCCAGGCGGTCCAGGAGCTTCCAGGAGGGCACTG 1328
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CGCAGGCCGGGAAGCAGTGCCCGCGCCAGGCGGTCCAGGAGCTTCCAGGAGGGCACTG 1320

Qy       1329 AGCGGCTGCTGCGGCGGCGGGACATCGGCAGGGACCACCAGGGGACGCTGAGGGGCGA 1388
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGCGGCTGCTGCGGCGGCGGGACATCGGCAGGGACCACCAGGGGACGCTGAGGGGCGA 1380

Qy       1389 TCTGTAAATAAACCTTTTTTCTTTTGTTTTTTT 1421
```

```
Db

RESULT 5
       |||||||||||||||||||||||||||||||||
  1381 TCTGTAAATAAACCTTTTTTTTTCTTTTGTTTTT 1413

ID   ADD19228 standard; cDNA; 1491 BP.
XX
AC   ADD19228;
XX
DT   15-JAN-2004 (first entry)
XX
DE   Human cDNA from secreted protein gene 45.
XX
KW   human secreted protein; cytostatic; antibacterial; virucide;
KW   cancer; bacterial infection; gastrointestinal-Gen; cardiant;
KW   muscular; reproductive; antiinflammatory; muscular-Gen;
KW   cardiovascular-Gen; nephrotropic; antiinflammatory; vasotropic;
KW   respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW   nootropic; antiallergic; cancer; bacterial infection; viral infection;
KW   neural disorder; immune system disorder; blood disorder;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   inflammatory disorders; proliferative disorder; human; ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO2003052377-A2.
XX
PD   26-JUN-2003.
XX
PF   06-NOV-2002; 2002WO-US035606.
XX
PR   07-NOV-2001; 2001US-0331046P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2003-533050/50.
DR   P-PSDB; ADD19303.
XX
PT   New isolated nucleic acids encoding signal transduction pathway component
PT   polypeptides, useful for diagnosing, treating, and/or preventing
PT   disorders, such as cancer, infections, cardiovascular and inflammatory
PT   diseases.
XX
PS   Claim 1; SEQ ID NO 55; 554pp; English.
XX

CC   The invention relates to an isolated nucleic acid molecule (cDNA)
CC   encoding a human secreted protein, representing one of 85 novel genes.
CC   Also included are recombinant vectors, host cells (expressing the
CC   protein), the secreted proteins (including their fragments, epitopes and
CC   homologues), an isolated antibody that binds specifically to the protein,
CC   diagnosing a pathological condition or susceptibility to a pathological
CC   condition (comprising determining the presence or absence of a mutation
CC   in the nucleic acid and diagnosing a condition based on the presence or
CC   absence of the mutation), diagnosing a pathological condition or
CC   susceptibility to a pathological condition (comprising determining the
CC   presence or amount of expression of the protein in a biological sample
CC   and diagnosing a condition based on the presence or amount of expression
CC   of the protein), preventing, treating or ameliorating a medical condition
CC   by administering the nucleic acid or protein to a mammalian subject,
CC   identifying a binding partner to the protein, the gene corresponding to
CC   the cDNA sequence, and identifying an activity in a biological assay
CC   (comprising expressing the nucleic acid in a cell, isolating the
CC   supernatant, detecting an activity in a biological assay and identifying
CC   the protein in the supernatant having the activity). The nucleic acids
CC   and proteins display the following activities Cytostatic, antibacterial,
CC   Virucide, Neuroprotective, Gynaecological, Gastrointestinal-Gen,
CC   Cardiant, Cardiovascular-Gen, Antiinflammatory, Muscular-
CC   Gen, Respiratory-Gen, Immunosuppressive, Cerebroprotective, Vasotropic,
CC   Nootropic, Antiallergic. The methods and compositions of the present
CC   invention are useful for diagnosing, treating, preventing and/or
CC   prognosticating disorders related to the novel polypeptides, such as
CC   cancer, bacterial or viral infections, and neural, immune system, blood,
CC   muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC   renal, inflammatory or proliferative disorders (many examples of these
CC   diseases and disorders are given in the specification). The present
CC   sequence encodes a novel secreted protein of the invention.
XX
SQ   Sequence 1491 BP; 293 A; 485 C; 410 G; 302 T; 0 U; 1 Other;

Query Match             99.4%;  Score 1412;  DB 10;  Length 1491;
Best Local Similarity   99.9%;  Pred. No. 0;
Matches 1412;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    8 GAAAGTAACGGCTTACAGACGAGTGAGAAATAGTTTCGGCTTCGGCCGGCGGTAGAAAACTCTGTC 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43 GAAAGTAACGGCTTACAGACGAGTGAGAAATAGTTTCGGCTTCGGCCGGCGGTAGAAAACTCTGTC 102

Qy   68 GGTACCAACCCCAGAGACGTTGAGGAGCAGCCCACCTTCCAGGCCTTCCTTAACGGAGGAGGTGC 127
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  103 GGTACCAACCCCAGAGACGTTGAGGAGCAGCCCACCTTCCAGGCCTTCCTTAACGGAGGAGGTGC 162

Qy  128 AGGACTCAGACTTCACCAGCCCACTCCGGTCCCAGCCGTTGTACGGCAAAGAGACGCCAAGGA 187
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  163 AGGACTCAGACTTCACCAGCCCACTCCGGTCCCAGCCGTTGTACGGCAAAGAGACGTCAAGGA 222

Qy  188 CGCGGCTCTCCCGCGGCGTCCAGGCAGCCCCAGCTTGCTGGCTTGGCCTTGGCCCCGGCCGCC 247
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  223 CGCGGCTCTCCCGCGGCGTCCAGGCAGCCCCAGCTTGCTGGCTTGGCCTTGGCCCCGGCCGCC 282

Qy  248 CACTCCGGCCGGCGGTGCAGCATGACCCGTGTGGAACGGCGTACGCCCTTTTTTACCCCCAGCC 307
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  283 CACTCCGGCCGGCGGTGCAGCATGACCCGTGTGGAACGGCGTACGCCCTTTTTTACCCCCAGCC 342

Qy  308 CCCGGCCATGCCGCAGGCGTTCCACGCGTTCCACTGCTCATCGTTATTCTCAGTGTGTTTCGGCTCT 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  343 CCCGGCCATGCCGCAGGCGTTCCACGCGTTCCACTGCTCATCGTTATTCTCAGTGTGTTTCGGCTCT 402

Qy  368 AGCAGCAGCAAGGTTCCTGCTCATCTTGCCCGGGGATCGGTTGGCCACTCGCGCGTGGTTTTGGTT 427
         ||||||| ||||||||||||||| ||||||||||||||| |||| ||||||||| |||| ||||
Db  403 AGCAGCAAGCTTCCTGCTCATTCTTGCCCGGGGATCCGGTTCGGTGGCCACTCGGCGCTCAGTGC 462

Qy  428 GGTGGAGAGTTCTTCTCAGTCGTGTTCATAGGCGCAGAAAATGTGGGCGTGCGCACTTCAGTGC 487
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  463 GGTGGAGAGTTCTTCTCAGTCGTGTTCATAGGCGCAGAAAATGTGGGCGTGCGCACTTCAGTGC 522

Qy  488 AGAATGGTTCGTGGGGTACAGTGAACACCAAACAATCCTACAAAAGCCTTCAGCGACGCG 547
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  523 AGAATGGTTCGTGGGGTACAGTGAACACCAAACAATCCTACAAAAGCCTTCAGCGACGCG 582

Qy  548 CGTTACAGCCCGGTCGTGTCGGGTCGTGCTCGGAGGGCCTCGAGGGCGCATTAATATTACACTCACAGG 607
         |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  583 CGTTACAGCCCGGTCGTGTCGGGTCTGCGTCGTGGGCCTCGAGGGCGCATTAATATTACACTCACAGG 642

Qy  608 GACCCCAGTGCATCAGCTGAACGAAGGACCCATTGACTACACGAGCAGTTCACCTGGCGTCT 667
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  643 GACCCCAGTGCATCAGCTGAACGAAGGACCCATTGACTACACGAGCAGTTCACCTGGCGTCT 702

Qy  668 GAAAGAGAATTACCCGCGCGGAGTACGTGGGAGAGGGGGGACCCAGT 727
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  703 GAAAGAGAATTACCGCGCGGAGTACGGGGAGCGCACTGGAGAAGGGGCTGCCGGACCCAGT 762

Qy  728 GCTTCTACCTGGGCGGAGAGTTCACACCGGAGTTAGCCCTTGCGGGTCGTACCAGTACCA 787
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  763 GCTCTACCTGGGCGGAGAGTTCACACCGGAGTTAGCCCTTGCGGGTCGTACCAGTACCA 822

Qy  788 CCTGGCGGGGACACTACGCCTCGGCCACGCTATGGGGTGGGGGGTTCTGCTTCTGGCTCCCTC 847
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  823 CCTGGCGGGGACACTACGCCTCGGCCACGCTATGGGGTGGGGGGTTCTGCTTCTGGCTCCCTC 882

Qy  848 CAAGGTGCTGCTCTCCACGCGCGGCCTCTCACGGAGGCGCCTGGCACTGCTGACCACCGG 907
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  883 CAAGGTGCTGCTCTCCACGCGCGGCCTCTCACGGAGGCGCCTGGCACTGCTGACCACCGG 942
```

```
RESULT 6
AAC95564
ID   AAC95564 standard; cDNA; 1460 BP.
XX
AC   AAC95564;
XX
DT   21-FEB-2001 (first entry)
XX
DE   Human secreted protein gene 44 SEQ ID NO:54.
XX
KW   Human; secreted protein; cytostatic; immunosuppressive; nootropic;
KW   neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
KW   antiinflammatory; vulnerary; antiulcer; anticonvulsant; antibacterial;
KW   antifungal; antiparasitic; cancer; immune disease; allergy;
KW   cardiovascular disorder; wound healing; infection; neurological disease;
XX
OS   Homo sapiens.
XX
FH
KW   88.
XX
PN   WO200061596-A1.
XX
PD   19-OCT-2000.
XX
PF   06-APR-2000; 2000WO-US008983.
XX
PR   09-APR-1999;       99US-0128703P.
PR   14-JAN-2000; 2000US-0176068P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Ruben SM, Komatsoulis G;
DR   WPI; 2000-611865/58.
DR   P-PSDB; AAB52055.
XX
```

PT   Fifty nucleic acid molecules encoding human secreted proteins, useful in
PT   the prevention, treatment and diagnosis of cancer, immune disorders,
PT   cardiovascular disorders and neurological diseases.
XX
PS   Claim 1; Page 443; 505pp; English.

```
XX
CC   Polynucleotide sequences AAC95521 - AAC95570 represent cDNA encoding 50
CC   human secreted proteins AAB52012 - AAB52161. Sequences AAB52062 -
CC   AAB52103 represent alternative polypeptides encoded by the genes, and
CC   amino acid sequences with which they share homology. The genes and
CC   proteins have activities dependent on the tissues and cells in which they
CC   are expressed. Examples of their activities include cytostatic;
CC   immunosuppressive; nootropic; neuroprotective; antiviral; antiallergic;
CC   hepatotropic; antidiabetic; antiinflammatory; antiulcer; vulnerary;
CC   anticonvulsant; antibacterial; antifungal; antiparasitic; and cardiant.
CC   The secreted proteins, polypeptides, antagonists and agonists may be
CC   useful in treating, preventing and/or diagnosing diseases and disorders
CC   such as cancer, particularly breast and ovarian cancer, and other cancers
CC   of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract,
CC   liver, lung, or urogenital. Immune disorders such as Addison's disease,
CC   allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC   diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC   arthritis and ulcerative colitis; cardiovascular disorders such as
CC   myocardial ischaemias; wound healing; neurological diseases such as
CC   cerebral anoxia and epilepsy; and infectious diseases such as viral,
CC   bacterial, fungal and parasitic infections may also be treated using the
CC   proteins and polynucleotides of the invention. Sequences AAC95512 -
CC   AAC95520 and AAB52011 are used in the isolation and characterisation of
CC   the proteins and polynucleotides of the invention
XX
SQ   Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;
```

```
Query Match      96.8%;  Score 1375.8; DB 3;  Length 1460;
Best Local Similarity   99.3%;  Pred. No. 0;
Matches 1413;  Conservative  0;  Mismatches  7;  Indels  3;  Gaps  3;

Qy      1 GAACCAGGAAAGTAACGGCTTACAGACAGTGAGAAATAGTTTCGCTGCCGGCTAGAAAAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     14 GAACCAGGAAAGTAACGGCTTACAGACAGTGAGAAATAGTTTCGCTGCCGGCTAGAAAAA 73

Qy     61 CTCTGTCGGTACCAACCCCAGAGACGTTGAGAGCAGCCCACCTTCCTTACGGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 CTCTGTCGGTACCAACCCCAGAGCGTTGAGAGCAGCCCACCTTCCTTACGGA 133

Qy    121 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCCTTGTTACGCAAAGAGACG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    134 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCCTTGTTACGCAAAGAGACG 193

Qy    181 CCAAGGACGGCCTCTCCGGGTCCAGGCAGCCCCACCTTGCTTGGCCTTGTTGCGCCGCTG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    194 TCAAGGACGGCGCCTCTCCGGGTCCAGGCAGCCCCAGCTTGCTTGGCCTTGTCCCCGCTG 253

Qy    241 CGTGCAGCACTCGGCCGCGGGTGCAGCAATGACCCGTGGAACGGGCGGGTACGCTTTTACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 CGTGCAGCACTCGGCCGCGGGTGCAGCAATGACCCGTGGAACGGGCTACTGCCTTTTTACC 313

Qy    301 CCCAGCCCCGGCCAGGCCTTCAGCGTTCCACTGCTTCCATCGTTATTCTTAGTGTTTT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    314 CCCAGCCCCGGCCATGCCGCAGGCTTCACTGCTTCCATCGTTATTCTTAGTGTTTT 373

Qy    361 TGGCTCTTAGCAGCAAGCTTCCATCTTGCCGGGGGATCGGTGGCCACTCGGCGCTGGT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374 TGGCTCTTAGCACCAAGCTTCCTGCTCAATCTTGCCGGGGATCGGTGGCCACTCGGCGCTGGT 433

Qy    421 TTTGGTTGGTGCAGGAGTTCTTCTTCAGTCTGTTTCATAGGCGCGAGAAAATTGTTGGCTGTGCACT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434 TTTGGTTGGTGCAGGAGTTCTTCTTCAGTCTGTTTCATAGGCGCAGAAAATTGTTGGCTGTGCACT 493

Qy    481 TCAGTGCAGAATGGTTCGTGGGGTACAGTGAACCACCAACACATCCTACAAAAGCCTTCAGCG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    494 TCAGTGCAGAATGGTTCGTGGGGTACAGTGAACCACCAACACATCCTACAAAAGCCTTCAGCG 553

Qy    541 CAGGCGCGCGTTACAGCCCGTGTCGGGTCTGCTGCGGGCGCATTAATATTACAC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    594 CAGGCGCGCGTTACAGCCCGTGTCGGGTCTGCTGCGGGCGCATTAATATTACAC 600
```

```
Qy    908 AGGCTTCGCGCTCTTCGCGGGTCTTCGCGGGTCTTCGCCTCCATCTCTAGGCGTGCCGCTCGCC 967
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    943 AGGCTTCGCGCTCTTCGCGGGTCTTCGCCTCCATCTCTAGGCGTGCCGCTCGCC 1002

Qy    968 GCTCCGCGCTAGGCTTCGCTCCGCGCTCACCACTCAGTACGGCCGCCTTCTGGGTCACGCT 1027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1003 GCTCCGCGCTAGGCTTCGCTCCGCGCTCACCACTCAGTACGGCCGCCTTCTGGGTCACGCT 1062

Qy   1028 GGCAACCGGCGTCCGTGCTTCCTCCGGAGCGGGCCGTTGGTGAGTCGTTCTTCAGTATGTTCG 1087
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1063 GGCAACCGGCGCGTCGTGCCTCTTCCTCCGGAGTCGTGGTGAGTCGTTCTTCAGTATGTTCG 1122

Qy   1088 GCCCAGGCGCTCTTCGCACCCTTCTGGGACCACTGCAGCCAAGGCCAAGGACTTGGCACCAGGAGG 1147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1123 GCCCAGGCGCTCTTCGCACCCTTCTGGGACTGCAGCCAAGGCCAAGGACTGCACCAGGAGG 1182

Qy   1148 GGGCTCACCTCTTATCTCGGCGGACCCCACTGCACAAGAGGCCGGCTCTCCCAGACTTAAA 1207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1183 GGGCTCACCTCTTATCTCGGCGGACCCACTGCACAAGAGGCCGGCTCTCCCAGACTTAAA 1242

Qy   1208 ATGTATCACCACTAACCTGTGAGGGGGGACCCAAATCTGGACTCCTTCCCGCCCTTGGGACA 1267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1243 ATGTATCACCACTAACCTGTGAGGGGGGACCCAAATCTGGACTCCTTCCCGCCCTTGGGACA 1302

Qy   1268 TCGCAGGCGGCGGAAGCAGTGCCCCCAGCCCTGGGGCCCAGGAGACTTCCAGGAAGGGCACT 1327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1303 TCGCAGGCGGCGGAAGCAGTGCCCCAGCCCTGGGGCCCAGGAGACTTCCAGGAAGGGCACT 1362

Qy   1328 GAGGCGCTGCTGTGGCCGCGAGGCCCTCGGACAATCCGCAGGCACCAGGGGAAAAGTTCTCCTGGGGCG 1387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1363 GAGGCGCTGCTGTGGCCGCGAGGCCCTCGGACAATCCGCAGGCACCAGGGGAAAAGTTCTCCTGGGGCG 1422

Qy   1388 ATCTGTAAAATAAAACCTTTTTTTCTTTTGTTTTTT 1421
          |||||||||||||||||||||||||||||||
Db   1423 ATCTGTAAATAAAACCTTTTTTTTTCTTTTGTTTTTT 1456.
```

```
RESULT 7

AB273640
ID   AB273640 standard; cDNA; 1460 BP.
XX
AC   AB273640;
XX
DT   12-MAY-2003  (first entry)
XX
DE   Secreted protein-encoding gene 360 cDNA clone HUFCL31, SEQ ID NO:370.
XX
KW   Human; secreted protein; cancer; tumour; hyperproliferative disorder;
KW   autoimmune disorder; inflammation; angiogenic diseases; AIDS;
KW   acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
KW   drug screening; chromosome identification; chromosome mapping;
KW   cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
KW   antianaemic; vulnerary; chromosome 9p21; gene; 88.
XX
OS   Homo sapiens.
XX
PN   WO200277013-A2.
XX
PD   03-OCT-2002.
XX
PF   26-MAR-2002; 2002WO-US009370.
XX
PR   27-MAR-2001; 2001US-0278650P.
PR   12-SEP-2001; 2001US-00950082.
PR   12-SEP-2001; 2001US-00950083.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2003-040578/03.
DR   P-PSDB; ABR01306.
XX
PT   New human secreted proteins and nucleic acids, useful for detecting or
PT   treating cancer or other hyperproliferative disorders, autoimmune
PT   disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
XX
PS   Claim 21; Page 1345; 2474pp; English.
XX
CC   AB273281-AB273697 represent cDNAs corresponding to 391 human secreted
CC   protein genes, and ABP00947-ABP01363 represent the proteins they encode.
CC   AB273698-ABZ74687 represent human secreted protein genomic fragments. The
CC   invention also encompasses antibodies specific for the secreted proteins,
CC   the use of the secreted proteins in drug screening and recombinant
CC   vectors and host cells comprising a nucleic acid of the invention. The
CC   secreted proteins are thought to be involved in biological activities
CC   associated with cellular signaling, cellular differentiation, cell
CC   migration, prohormone activation and neurotransmitter activity. The
CC   secreted proteins, nucleic acids encoding them, antibodies or antibody
CC   fragments specific for the secreted proteins, and modulators of protein
CC   activity are useful for diagnosing or treating cancers or other
CC   hyperproliferative disorders. Additionally, the secreted proteins and
CC   their nucleic acids may also be used in the treatment of autoimmune
CC   disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
CC   (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
CC   wound healing. Nucleic acids of the invention may be used for chromosome
CC   identification, chromosome mapping, in gene therapy, for identifying
CC   individuals from minute biological samples, as hybridisation probes, and
CC   as molecular weight markers. The present sequence represents a human
CC   secreted protein-encoding cDNA clone of the invention
XX
SQ   Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;

Query Match             96.8%;  Score 1375.8;  DB 8;  Length 1460;
Best Local Similarity   99.3%;  Pred. No. 0;
Matches 1413;  Conservative    0;  Mismatches    7;  Indels    3;  Gaps    3;

Qy      1 GAACCAGAGAAAGTAACGGCTACAGACAGTGAGAGAAATAGTTTCGCTCGCCGCGGGCTAGAAAAA 60
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db     14 GAACCAGAGAAAGTAACGCGTACAGACAGTGAGAGAAATAGTTTCGCTCGCCGCGGGCTAGAAAAA 73

Qy     61 CTCTGTCGGTACCAAACCCCAGAGACGTTGAGAGCAGCCCACCTTCCACGCTTCCTTAACGGA 120
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     74 CTCTGTCGGTACCAAACCCCAGAGAGGTTGAGAGCAGCCCACCTTCCACGCTTCCTTAACGGA 133

Qy    121 GAGGTGCAGGACTCAGATCTTCACCGAGCCCACTCGGTCCCAGCCTTGTACGCAAAGAGACG 180
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    134 GAGGTGCAGGACTTCACCGAGCCCACTCGGTCCCAGCCTTGTACGCAAAGAGACG 193

Qy    181 CCAAGGACGCGCTTCTCCCGCGTCCCAGGCAGCCCCAGCTTGGCTTGGCTTGCCTGCCGCCTG 240
          ||||||||||||||  ||||| ||||||||||||||||||||| ||||||||||||||||||
Db    194 TCAAGGACGCGCTTCCCGCGGTTCCAGGCAGCCCCAGCTTGCTGCTTGCCTGCCGCCTG 253
```

(Alignment continues: query/database positions from 554/601 through 1435/1421, dense nucleotide sequence with identity bars.)

```
Qy   241 CGTGCAGCAACTCGGCCGTGCGAGCAATGACCCTGTGGAAACGGCGTACTGCCTTTTTACC  300
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   254 CGGTGCAGCAGCACTCGGCGCCGGCGGTGCAGCAATGACCCTGTGGGAACGGCGTACTGCCTTTTTACC  313

Qy   301 CCCAGCCCCGGCATGCCCACGCGGCTTCACGGTTCCACTGTCTCATCGGTTATTCTAGTGTTTT  360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   314 CCCAGCCCCGGCATGCCCAGGCGGCTTCACGGTTCCACTCGCTCATCGGTTATTCTAGTGTTTT  373

Qy   361 TGGCTCTPAGCAGCAAGCTTCCTGTCTCATCTTGCCGGGGATCCGGTGGGCCACTCGGCGTGT  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   374 TGGCTCTPAGCAGCGAAGCTTCCTGCTCATCTTGCCGGGGATCCGGTGGGCCACTCGGCTGGT  433

Qy   421 TTTGGTTGGTGAGAGTTCTTCTCTCAGTGTTCGTTAGTTCATPAGGGGCAGAAATTGTTGGCTGTGCACT  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   434 TTTGGTTGGTGAGAGTTCTTCTCTCAGTGTTCGTTCATPAGGGGCAGAAATTGTTGGGCGTGCACT  493

Qy   481 TCAGTGCAGAATGGTTCGTTCGTTAGTAGAACCAACACATCCTTACAAAGCCTTCAGCG  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   494 TCAGTGCAGAATGGTTCGTTCGTTAGTAGAACCAACACATCCTTACAAAGCCTTCAGCG  553

Qy   541 CAGGCGCGCCTTACAGCCCGTGTCGTCTGCCTTCGTCGTGGGGCCTTGGGCATTAATATTACAC  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   554 CAGGCGCGCCGTTACAGCCCGTGTCGGTCTGCCTTCGTCGTGGGGCCTGGGGCATTAATATTACAC  613

Qy   601 TCACAGGGACCCCAGTGCATCAGCTGAACGAGACACCATTGACTACTACAGAGCCAGTTCACTT  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   614 TCACAGGGACCCCAGTGCATCAGCTGAACGAGACACCATTGACTACTACAGAGCCAGTTCACTT  673

Qy   661 GGGGTCTGAAAGAGAATTACGCGCGGGAGTACGCCACTGGAGAAGGGGGGTGCCGG  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   674 GGGGTCTGAAAGAGAATTACGCCCCGGGAGTACGCCACTGGAGAAGGGGGGTGCCGG  733

Qy   721 ACCCAGTGCTCTACCTGGCGGAGAAGTTCACACCGAGTAGCCCTGTGCGGCCTGTACCACC  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   734 ACCCAGTGCTTCACCTGGCGGAGAAGTTCACACCGAGTAGCCCTTGCGCGGCCTGTACCACC  793

Qy   781 AGTACCACCTGGGCGGGACACTACGCCTCGGCCACGGCTGGGCGTTCTGCTTCTGGC  840
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   794 AGTACCACCTGGGCGGGACACTACGCCTCGGCCTATGGGTGGGCGTTCTGCTTCTGGC  853

Qy   841 TCCTCTCCAACGTGCTCTGCTTCCTCCACGGCGCGCCTCACCTTCTACGGAGGGCACTGCTGA  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   854 TCCTCTCCAACGTGCTCTGCTTCCTCCACGGCCCCGGCCTCTAGGAGGGGCCACTGCTGA  913

Qy   901 CCACCGGACGCCTTCGGCGCTTCTGGGGTCTTCGGCGCTCCATTCTCTAGCCGTGCCGC  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   914 CCACCGGAGCCTTCGGCGCTTCTGGGGTCTTCGGCGCTCCATTCTCTAGCCGTGCCGC  973

Qy   961 TCTGCCCCGCTCCGCCTCCGGGCTCTCCCGGCTCACCAGTCAGTAGGCCCGCGCCGCCTTCTCGGG 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   974 TCTGCCCCGCTCCGCCTCCAGGCTCTCCCGGCTCACCAGTCAGTAGGCCCGGCGCGCCTTCTCTGGG 1033

Qy  1021 TCAGGCTGGCAACCGGCGGTCCTGTGCCGTGGTGAGTGCGTCCAGT 1080
             ||||||||||||||||||||||||||||||||||||||||||||||
Db  1034 TCACGCCTGGCAACCGGCGGTCCTGTGCCTCTCCCAGGCCGTGGTGAGTCTCCAGT 1093

Qy  1081 ATGTTCGGCGCCAGCCGCTCTCGCACCCTCTGGACCAAAGGCACTGC-AGCCAG 1139
             ||||||||||||||||||||||||||||||||||||||||||||||||  ||||
Db  1094 ATGTTCGGCGCCAGCCGCTCTTCGACACCCTTCTGGACCAAAAGCGGACTGCAAAGCCAG 1153

Qy  1140 GAGAGAGGGGGGCCTCACCTCTTATCCTCGGGGAACCGACCAGCCGGCGCTCTCCCCA 1199
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1154 AAGAGGGGGGGCTCACCTCTTATCCT-GGGGACCCACTGCACAACGACGCGGTTTTCCCA 1212

Qy  1200 GAC-TTAAAATGTATCACCACTAACCTGTGAGGGGGGGACCCAATCTGGACTCCTTCCCCGC 1258
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1213 GACTTTAAAATGTATCACCACTAACCTGTGAGGGGGGGACCCAATCTGGACTCCTTCCCCGC 1272

Qy  1259 CCTGGGACATCGCAGGCCCGAAGCAGTGCCCGCCCAGGCCTGGGCGCCAGGAGAGCTCCAGG 1318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1273 CTTGGGACATCGCAGGCCGGGGAAAGCAGTGCGCGGCCTGGGCCTGGGCCAGGGAGAGCTCCAGG 1332

Qy  1319 AAGGGCACTGAGCGGCCTGCTCGGCGCGCGCAGCCACCAGGGACGGGAAAGTCT 1378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1333 AAGGGCACTGAGCGGCCTGCTCGGCGCGCGCAGCCACCAGGGACGGGAAAGTCT 1392
```

```
Db  1333 AAGGGCACTGAGCGGCCTGCTCGGCGCGCGCAGCCACCAGGGACGGGAAAGTCT 1392
            ||||||||||||||||||||||||||||||||||||||||||||||
Qy  1379 CCTGGGGGCCGATTCGTGTAAATAAAAACCTTTTTTTCTTTTTTGTTTTT 1421
            ||||||||||||||||||||||||||||||||||||||||||||||
Db  1393 CCTGGGGGCCGGATTCGTGTAAATAAACCTTTTTTTTCTTTTTTTGTTTTT 1435
```

RESULT 8

```
ADC20289
ID   ADC20289 standard; DNA; 1460 BP.
XX
AC   ADC20289;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human secreted protein coding sequence #228.
XX
KW   gene therapy; human; secreted protein; haemopoietic disorder;
KW   haematological disorder; anaemia; haemophilia; inflammatory disorder;
KW   inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
KW   leukaemia; wound healing; epithelial cell proliferation disorder;
KW   immune disorder; autoimmune disorder; asthmatic disorder;
KW   cardiovascular disorder; atherosclerosis; myocarditis;
KW   infectious disease; HIV; AIDS; endocrine disorder; diabetes;
KW   gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.
OS   Homo sapiens.
XX
PN   WO20092787-A2.
XX
PI   Rosen CA,  Ruben SM;
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PR   12-SEP-2001; 2001US-00950083.
PR   12-SEP-2001; 2001US-00950082.
PR   27-MAR-2001; 2001US-02786509P.
XX
PT   New human secreted proteins and nucleic acid molecules, useful for
PT   preparing a diagnostic or pharmaceutical composition for diagnosing,
PT   preventing or treating hematopoietic or hematologic disorder, e.g.
PT   anemia or hemophilia.
XX
PF   26-MAR-2002; 2002WO-US009257.
XX
PD   21-NOV-2002.
XX
DR   WPI; 2003-129287/12.
XX
PS   Claim 1; SEQ ID NO 238; 1512pp; English.
XX
CC   The invention comprises the amino acid and coding sequences of human
CC   secreted proteins. The DNA and protein sequences of the invention are
CC   useful for detecting, preventing, diagnosing, prognosticating, treating
CC   or ameliorating: haematopoietic or haematological disorders (e.g. anaemia
CC   and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease
CC   and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia);
CC   wound healing and disorders of epithelial cell proliferation; immune
CC   disorders (e.g. autoimmune disorders and asthmatic disorders);
CC   cardiovascular disorders (e.g. atherosclerosis and myocarditis);
CC   infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
CC   and gastrointestinal disorders (e.g. duodenal ulcers and
CC   gastroenteritis). The present DNA sequence encodes a human secreted
CC   protein of the invention.
XX
SQ   Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;
```

```
Query Match          96.8%;  Score 1375.8;  DB 10;  Length 1460;
Best Local Similarity  99.3%;
Matches 1413;  Conservative  0;  Mismatches  7;  Indels  3;  Gaps  3;

Qy     1 GAACCAGGAAAGTAACGGCTACAGACAGTGAGAAATAGTTCGTCGCGGCCTAGAAAAA 60
```

```
Db     14  GAACCACGGAAAAGTAACCGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGCGGCTAGAAAA  73
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy     61  CTCTGTCGGTACCAAACCCCAGGACGGTTGAGAGAGCAGCCGCAGCGTTCCTTAACGGA  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74  CTCTGTCGGTACCAAACCCCAGAGAGCGAGCGCCACCTTCCTTAACGGA  133

Qy    121  GAGGTGCAGGACTCAGACTTCACCAGAGCGGTCCAGCCACTCGGTCAAAGAGACG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    134  GAGGTGCAGGACTCAGACTTCACCAGAGCGGTCCAGCCACTCGGTCAAAGAGACG  193

Qy    181  CCAAGGACGACGCGCTCTCCGCGTCCAGGCCAGCGCCCCAGCGTTGCTGGCTTGTGCCGCCTG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    194  TCAAGGACGACGCGCTCTCCCGGTCCAGCGCCCCAGCGTTGCTGGCTTGTGCCGCCTG  253

Qy    241  CGTGCAGCACTCGGCCGCCGGTGCAGCAATGACCCTGTGGAAACGGTACTGCCTTGCTGGCCTG  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254  CGTGCAGCACTCGGCCGCCGGTGCAGCAATGACCCTGTGGAAACGGTACTGCCTTGCTGGCCTG  313

Qy    301  CCCAGCCCCGGCCGCCAGGCTTCAGGGTTCCACTGCGTTCCATCGTTAATTCTAGTGTTTT  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    314  CCCAGCCCCGGCCGCCAGGCTTCACGGTTCCACTGCGTTATTCTAGTGTTTT  373

Qy    361  TGGCTCTTAGCAGCAAGCTTCCTGCTCATCTTGCCGGGGAGATCCGGTGGCCACTCGCGGCTGGT  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374  TGGCTCTTAGCAGCAAGCTTCCTGCTCATCTTGCCGGGGATCCGGTGGCCACTCGCGGCTGGT  433

Qy    421  TTTGGTTGGTGAGAGTTCTTCTTCCAGTGTTCATAGGGGCAGAAAATTGTGGCTGTGTGCACT  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434  TTTGGTTGGTGAGAGTTCTTCTTCCAGTGTTCATAGGGGCAGAAAATTGTGGCTGTGTGCACT  493

Qy    481  TCAGTGCAGAATGGTTCGTTCGTTGGGGTACAGTGAACACACAACATCCTACACAAGCCTTCAGCG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    494  TCAGTGCAGAATGGTTCGTTCGTTGGGGTACAGTGAACACCAACATCCTACAAAGCCTTCAGCG  553

Qy    541  CAGGCGCGGTTACAGCCCGTGTCGGTCTTGCTCGTCGTGGCGGTCCATTAATATTACTACAC  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    554  CAGGCGCGGTTACAGCCCGTGTCGGTCTTGCTCGTCGTGGCGGTCATTAATATTACAC  613

Qy    601  TCACAGGGACCCCAGTGCATCAGCTGAACGAGAGACCAGCAGTGCACT  660
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    614  TCACAGGGACCCCAGTGCATCAGCTGAACGAGACCAATTGACTACTACAACGAGCAGTTCACCT  673

Qy    661  GGCGTCTGAAAGAGAATTACGCGCGGAGTACGCGGAAACGGCCGCTGCCGG  720
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    674  GGGCGTCTGAAAAGAGAATTACGCCCGCGGAGTACGCACTGGAGAAGGGCACTGGCGGG  733

Qy    721  ACCCAGTGCTCTACCTGGGCGGAGAAGTTCACACACGGAGTAGCCCTTGCGGCCGTGTACCACC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    734  ACCCAGTGCTCTACCTGGGCGGAGAAGTTCACACACGAGTAGCCCTTGCGCGCGTGTACCACC  793

Qy    781  AGTACCACCTGGGCGGGGACACTAGGCCTCGGCGGCCCTTCTGGG  840
           ||||||||||||||||||||||||||||||||||||||||||||
Db    794  AGTACCACCTGGGCGGGGACACTAGGCCTCGGCGGCCTTCTGGG  853

Qy    841  TCCTCTCCAACGTGCTGTGCTCTCACGCCGCCGCCTGGCACTGCTGA  900
           |||||||||||||||||||||||||||||||||||||||||||||
Db    854  TCCTCTCCAAAGTGCTGCGTCTCCAACGCGCGGCCTGGCCACTGCTGA  913

Qy    901  CCACCGGAGCCTTTCGCGCGTTCTTCGCGCGTTCGCGCGTTCTCCAATCTTCTAGCGTGCGCGC  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    914  CCACCGGAGCCTTCGCGCGTTCTTCGGGGGTTCGCCCTTCCATCTCTAGCGTGCGCG  973

Qy    961  TCTTGCCCCGCGCTGGCGGGGGACACTAGGCTACCTCCAGTCAGTACGCCCTTCTGGG  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    974  TCTGGCCCCGCTCCGGCTCACCACTCTCCGGCTCAGGTACGCCGCCTTCTCGGG  1033

Qy   1021  TCACGCTGGCCAACCGGCCGTCCTGTGCTCTTCGCGGAGGGGGCGGTGGTGAGGTCTTCAGT  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1034  TCACGCTGGCAACCGGCGGTCCTGTTGCTCTTCCCCAGAGGGGGCCGGTGGTGAGTCTTCAGT  1093

Qy   1081  ATGTTCGGCCCAGGCGTCTTCGCACCCTTCTGGACCAAAGGCCCAAGGGCACTGACTGCAAGCCAG  1139
           ||||||||||||||||||||||||||||||||||| ||||  ||||
Db   1094  ATGTTCGGCCCAGCGTCTTCGCACCCTTCTGGACCAAAGGCCAAGGGACTGCAAGCCAG  1153
```

```
Qy   1140  GAGAGGAGGGGGCTCACCTTCTTATCCTCGGCGACGCCCACTGCACAAGCCAGGCGCTCCCA  1199
           ||||||||||||||||||||||||||||||||||||
Db   1154  AAGAGGGGGGGCTCACCTTCTTATCCT-GGGGACCCCACTGCACAAGCACGGCGCGCTTCCCA  1212

Qy   1200  GAC-TTAAAAATGTATCACCACTAACCTGTGAGGGGGGACCCAATGCTGGACTCCTTCCCGC  1258
           |||  ||||||||||||||||||||||||||||||||||||||||||||||
Db   1213  GACTTTAAAAATGTATCACCACTAACCTGTGAGGGGGACCCAAATCTGGACTCCTTCCCGC  1272

Qy   1259  CTTCGGGGACAATCGCAGGCCCGGGGAAGCCAGTGTCCCGGCCAGCCTTCCAGG  1318
           ||||||||||||||||||||||||||||||||||||||||||
Db   1273  CTTCGGGGACAATCGCAGCCCCGGGGAGCCAGTGCCGGGCCAGCCTTCCAGG  1332

Qy   1319  AAGGGCACTGAGCCGCTGCTGGGCGCGGAGCACATCCGGCACGTCGGCCCGGGAAAGTCT  1378
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   1333  AAGGGCACTGAGCCGCTGCTGGGCGCGGAGCACATCGGCACGTCCGGCCAGGGAAAGTCT  1392

Qy   1379  CCTGGGGGCGGATCCTGTAAATAAACCTTTTTTTTGTTTTTT  1421
           ||||||||||||||||||||||||||||||||||||||||
Db   1393  CCTGGGGGCGGATCCTGTAAATAAACCTTTTTTTTGTTTTTT  1435
```

RESULT 9
AB267235
ID   AB267235 standard; cDNA; 1460 BP.
XX
AC   AB267235;
XX
DT   26-MAR-2003 (first entry)
XX
DE   Human secreted protein encoding cDNA SEQ ID NO 355.
XX
KW   Human; secreted protein; nootropic; neuroprotective; cytostatic;
KW   virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW   vulnerary; antibacterial; antiparkinsonian; antisickling; antianaemic;
KW   antiarthritic; anticancer; antirheumatic; hepatotropic; cerebroprotective;
KW   antiinflammatory; antiallergic; antidiabetic; antiulcer; anticonvulsant;
KW   antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW   cardiovascular disorder; neurological disease; nephrotropic;
KW   gene therapy; gene; chromosome 9p21; ds.
OS   Homo sapiens.
XX
PN   WO20027118G-A2.
XX
PD   03-OCT-2002.
XX
PF   26-MAR-2002; 2002WO-US009188.
XX
PR   27-MAR-2001; 2001US-0278650P.
PR   12-SEP-2001; 2001US-00950082.
PR   12-SEP-2001; 2001US-00950083.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2003-040583/03.
DR   P-PSDB; ABP99814.
XX
PT   New human secreted proteins encoded by genes contained in cDNA clones
PT   (e.g. HGCAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT   multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT   West Nile fever.
XX
PS   Claim 7; Page 1353; 2423pp; English.
CC
CC   The invention relates to novel human genes (ABZ66891-ABZ68209) and the
CC   encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC   treating or ameliorating medical conditions e.g. by protein or gene
CC   therapy. The genes are isolated from a range of human tissues disclosed
CC   in the specification. The nucleic acids, proteins, antibodies and

```
CC   (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC   (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC   adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC   lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC   allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC   diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC   arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC   myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC   cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC   bacterial, fungal and parasitic infections
XX
SQ   Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;

Query Match             96.8%;  Score 1375.8;  DB 10;  Length 1460;
Best Local Similarity   99.3%;
Matches 1413;  Conservative    0;  Mismatches    7;  Indels    3;  Gaps    3;
```

```
Qy     1 GAACCAGAGAAAGTAACGGCTACACAGACAGTAGAAATAGTTTCGCTCGCCGGCCTAGAAAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    14 GAACCAGAGAAAGTAACGGCTACACAGACAGTAGAAATAGTTTCGCTCGCCGGCCTAGAAAAA 73

Qy    61 CTCTGTCGGTACCAACCCCAGAGAGCGGTTGAGAGCAGCCCACCTCCACCGCTTCCTTAACCGGA 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    74 CTCTGTCGGTACCAACCCCAGAGAGCGGTTGAGAGCAGCCCACCTCCACCGCTTCCTTAACCGGA 133

Qy   121 GAGGTGCAGGACTTCAGACTTCACCAGCCCACTCGGTCCCAGCCCTTGTACGGCAAAGAGACG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   134 GAGGTGCAGGACTTCAGACTTCACCAGCCCACTCGGTCCCAGCCCTTGTACGGCAAAGAGACG 193

Qy   181 CCAAGGACGCCGCTCTTCCCCGTCAGCCAGCCCACTCGGTCCCAGCCCTTGCTGGCTTGCCTGCCGCCTG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   194 TCAAGGACCGCGCTCTCCCCGCGTCCCAGCCAGCCGCTTGCTGGCTTGCCTGCCGCCTG 253

Qy   241 CGTGCAGCACTCCGCCGCCGGGTGCAGCATGACCCTGTTGTGGACGGGTACTCCCTTTTTTACC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   254 CGTGCAGCACTCCGCCGCCGGGTGCAGCATGACCCTGTTGTGGACGGGTACTCCCTTTTTTACC 313

Qy   301 CCCAGCCCCGGGCCATGCCGCCGGGCTTCAGCCGTTCCACTGCTTCATCGGTTATTCTAGTGTTTT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   314 CCCAGCCCCGGGCCATGCCGCCGGGCTTCAGCCGTTCCACTGCTTCATCGGTTATTCTAGTGTTTT 373

Qy   361 TGGCTCTAGCAGCAAGCTTCCTGCTCATCTTGCCCGGGACCGTGGCCACTCGCCGCTGGT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   374 TGGCTCTAGCAGCAAGCTTCCTGCTCATCTTGCCCGGGACCGTGGCCACTCGCCGCTGGT 433

Qy   421 TTTGGTTGGTGAGAGTTCTTCTCAGTCTGTTCAATAGGCGCGAGAAATTGTGGCTGTGCACT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   434 TTTGGTTGGTGAGAGTTCTTCTCAGTCTGTTCAATAGGCGCGAGAAATTGTGGCTGTGCACT 493

Qy   481 TCAGTGCAGAATGGTTCGTGGGTACAGTGAACACCAAACACTCTACAAAGCCTTCAGCG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   494 TCAGTGCAGAATGGTTCGTGGGTACAGTGAACACCAAACACTCTACAAAGCCTTCAGCG 553

Qy   541 CAGGCGCGGTTACAGCCCGTGTCTCGTTCGTGGGCCTTCGTGGCCGTCGCCATTAATAATTACAC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   554 CAGGCGCGGTTACAGCCCGTGTCTCGTTCGTGGGCCTTCGTGGCCGTCGCCATTAATAATTACAC 613

Qy   601 TCACAGGGACCCAGTGCATCAGCTGAACGAGGACCATTGACTACTACAACGGAGCAGTTCACCT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   614 TCACAGGGACCCAGTGCATCAGCTGAACGAGGACCATTGACTACTACAACGGAGCAGTTCACCT 673

Qy   661 GGCGTCTGAAAGAGAATTACGGCCGGGAGTACGCGGGGAGAAGGGCCTGCCGG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   674 GGCGTCTGAAAGAGAATTACGGCCGGGAGTACGCGGGGAGAAGGGCCTGCCGG 733

Qy   721 ACCCAGTGCTCTACCTGGGCGGAGAGTTCACACCGAGGAGAGCGGCTGCCACC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   734 ACCCAGTGCTCTACCTGGGCGGAGAGTTCACACCGAGGAGAGCGGCTGTACCACC 793

Qy   781 AGTACCACCTGGGCGGGACACTACGCCTCGGCCCACCGCTATGGGTAGGTGGGCGTTCTGCTTCTCGGC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   794 AGTACCACCTGGGCGGGACACTACGCCTCGGCCCACCGCTATGGGTAGGTGGGCGTTCTGCTTCTCGGC 853
```

```
Qy   841 TCCTCTCCAACGTGCTGCCTCTCCAACGCGCCGCGCCGCCGGCTTCTTCTACGGAGGCCTGCCACTGCTGA 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   854 TCCTCTCCAACGTGCTGCTTCCTCCAACGCGCCCGCGCCGCCTTCTTCTACGGAGGCCTGCCACTGCTGA 913

Qy   901 CCACCGGAGCCGCTTGCGCGCTCTTCGGGGTCTTCGGCCTCCATCTCTACGGTGCCGC 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   914 CCACCGGAGCCGCTTGCGCGCTCTTCGGGGTCTTCGGCCTCCATCTCTACGGTGCCGC 973

Qy   961 TCTGCCCGCTCCGCCTAGGCTCTCCGCGCGCTCACCACTCAGTATACGGGCCCCTTCTCGGG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   974 TCTGCCCGCTCCGCCTAGGCTCTCCGCGCGCTCACCACTCAGTATACGGGCCCCTTCTCGGG 1033

Qy  1021 TCACGGCTGGGCAACCCGGCGTGCCTCTTCCTCGGAGAGGGGGCCTGGGTGAGTCTTCCAGT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1034 TCACGGCTGGGCAACCCGGCGTGCCTCTTCCTCCCAGGGGCCTGGGTGAGTCTTCCAGT 1093

Qy  1081 ATGTTTCGGCCCCAGGCCGTTCTTCGGCACCCTTCTGGACACAAACGCCAAAGGACGACTGGC-AGCCAG 1139
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1094 ATGTTTCGGCCCCAGGCCGTTCTTCGGCACCCTTCTGGACACAAACGCCAAAGGACGACTGGCAGCCAG 1153

Qy  1140 GAGAGAGGGGGCTCACCTCTTATCCTCGGGGACCCACTGCACAAGCAGCCGCGCTCCCA 1199
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1154 AAGAGGGGGGGCTCACCTCTTATCCT-GGGGACCCACTGCACAAGCAGCCGCGCTTCCCA 1212

Qy  1200 GAC-TTAAAATGTATCACACTACCTGTCAGGAGACCAAATCTGGACTCCTTCCCCGC 1258
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1213 GACTTTAAAAATGTATCACACTAACCTGTCAGGAGACCAAATCTGGACTCCTTCCCCGC 1272

Qy  1259 CTTGGGACATCGCACGGCCGGAAGCAGTGCCGGGCCAGGAGAGCTCCAGG 1318
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1273 CTTGGGACATCGCACGGCCGGAAGCAGTGCCGGGCCAGGAGAGCTCCAGG 1332

Qy  1319 AAGGGGCACTGAGGCGCTGGGCGCGGAGGGCGGCATCCGCAGGGACCAGGGAAAGTCT 1378
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1333 AAGGGGCACTGAGGCGCTGGGCGCGGAGGGCGGCATCCGCAGGGACCAGGGAAAGTCT 1392

Qy  1379 CCTGGGGGCGGATCTGTAAATAAAACCTTTTTTTCTTTTGTTTTTT 1421
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1393 CCTGGGGGCGGATCTGTAAATAAAACCTTTTTTTTCTTTTGTTTTTT 1435
```

```
RESULT 10
ID   ADU01671  standard; cDNA; 960 BP.
XX
AC   ADU01671;
XX
DT   27-JAN-2005  (first entry)
XX
DE   Novel human polynucleotide seqid 138.
XX
KW   cytostatic; antipsoriatic; antiinflammatory; gene therapy; Nanodisc;
KW   proliferative disorder; inflammatory disorder; immune disorder;
KW   metabolic disorder; bone disorder; CNS disorder; cancer; psoriasis;
KW   ulcerative colitis; human; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004093804-A2.
XX
PD   04-NOV-2004.
XX
PF   19-APR-2004; 2004WO-US012047.
XX
PR   18-APR-2003; 2003US-0463708P.
PR   18-APR-2003; 2003US-0463732P.
PR   02-MAY-2003; 2003US-0467199P.
PR   02-MAY-2003; 2003US-0467230P.
PR   19-MAY-2003; 2003US-0471306P.
PR   19-MAY-2003; 2003US-0471336P.
PR   08-JUL-2003; 2003US-0485223P.
PR   08-JUL-2003; 2003US-0485224P.
PR   14-JUL-2003; 2003US-0486446P.
```

us-09-937-059-57.rng

```
PR   14-JUL-2003;  2003US-0486480P.
PR   08-AUG-2003;  2003US-0493573P.
PR   08-AUG-2003;  2003US-0493577P.
PR   08-SEP-2003;  2003US-0505059P.
XX
PA   (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI   Lee E, Hestir K, Chu K, Masuoka L, Williams LT;
XX
DR   WPI; 2004-775861/76.
DR   P-PSDb; AD002403.
XX
PT   New first nucleic acid molecule comprising a polynucleotide sequence
PT   given in the specification, useful in preparing a composition for
PT   diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
XX
PS   Claim 1; SEQ ID NO 138; 291pp; English.
XX
CC   The invention describes a new first nucleic acid molecule comprising a
CC   polynucleotide sequence given in the specification. Also described are:
CC   an animal injected with the nucleic acid molecule; a second nucleic acid
CC   molecule comprising a second polynucleotide sequence that is at least
CC   about 70, 80, 90 or 95% homologous to the first nucleic acid molecule or
CC   that hybridises to the first polynucleotide sequence under high
CC   stringency conditions; a vector comprising the nucleic acid molecule and
CC   a promoter that drives the expression of the nucleic acid molecule; a
CC   host cell transformed, transfected, transduced or infected with the
CC   nucleic acid molecule; a nucleic acid composition comprising a carrier or
CC   buffer and one or more compositions comprising the nucleic acid
CC   molecule, vector or host cell; a substantially purified polypeptide; an
CC   animal injected with the polypeptide; a polypeptide composition
CC   comprising the polypeptide molecule and a carrier or buffer; a cell
CC   culture medium comprising the polypeptide or transfected cells
CC   transfected with the polynucleotide; making a transformed, transfected,
CC   transduced, or infected host cell; synthesising Nanodiscs simultaneously
CC   and for synthesising a series of simultaneously-synthesised Nanodiscs
CC   sequentially utilising a dynamic system; preparing a hydrophobic protein
CC   for determination of crystal structure; immunising a non-human animal;
CC   screening for modulators of hydrophobic protein activity; a diagnostic
CC   kit; determining the presence of the nucleic acid molecule or its
CC   complement; determining the presence of an antibody to the polypeptide in
CC   a sample; an antibody specifically recognising, binding to or modulating
CC   the biological activity of at least one polypeptide encoded by a nucleic
CC   acid molecule or its biologically active fragment; an antibody
CC   composition comprising the antibody and a carrier; a bacteriophage, where
CC   the antibody is displayed on the bacteriophage; a bacterial cell
CC   comprising the bacteriophage; a non-human animal injected with the
CC   antibody composition; a host cell that secretes the antibody; making an
CC   antibody; diagnosing a disease, disorder, syndrome, or condition
CC   comprising cancer, or proliferative, inflammatory, immune, metabolic,
CC   bone, CNS, genetic, bacterial and viral diseases, disorders, syndromes or
CC   conditions in a patient; prophylactic or therapeutic treatment of a
CC   subject; an isolated modified cell comprising at least one first
CC   heterologous nucleic acid molecule, where the first heterologous nucleic
CC   acid molecule comprises a first polynucleotide sequence that encodes a
CC   first polypeptide; a non-human animal deficient in the polypeptide or
CC   that over-expresses the polypeptide; isolated tissues derived from the
CC   non-human animal; and one or more cells derived from the non-human
CC   animal. The nucleic acid is useful in preparing a composition for
CC   diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
CC   This sequence encodes a novel human polypeptide of the invention.
XX
SQ   Sequence 960 BP; 168 A; 321 C; 260 G; 211 T; 0 U; 0 Other;

Query Match             67.4%;  Score 958.4;  DB 13;  Length 960;
Best Local Similarity   99.9%;  Pred. No. 1.8e-229;
Matches  959;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy          267 ATGACCCTGTGGAACGGGTACTGCCTTTTTACCCCCAGCCCCGCCCGGGCATGCCGCAGGCTTC 326
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 ATGACCCTGTGGAACGGGTACTGCCTTTTTACCCCCAGCCCCGGGCATGCCGCAGGCTTC 60
```

```
Qy          327 AGCGGTTCCACTGCTCATCGTTATTCGTTATTCTAGTGTTTTTTGGCTCTAGCAGCAAAGCTTCCTGCTC 386
                ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db           61 AGCGGTTCCACTGCTCATCGTTATTCGTTATTCTAGTGTTTTTTGGCTCTAGCAGCAGAAAGCTTCCTGCTC 120

Qy          387 ATCTTGCCGGGGATCCGTTGGGCCACTCCGCGGCTGGTTGTTGTTTGGTTGGTTGGAGGTTCTTCTCAGT 446
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 ATCTTGCCGGGGATCCGTTGGGCCACTCCGCGGCTGGTTGTTTGGTTGGTTGGAGGTTCTTCTCAGT 180

Qy          447 CTGTTCATPAGGCGGCAGAAATTGTGGCTGTGCACTTCAGTGCAGAAATGGTTCGTTCGTTGGGGTTACA 506
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 CTGTTCATPAGGCGGCAGAAATTGTGGCTGTGCACTTCAGTGCAGAAATGGTTCGTTCGTGGGGTTACA 240

Qy          507 GTGAACACCAACACATCCTACAAAAGCCTTCAGCGAGCGAGCGGCGGTTACAGCCCCGGTTCGGT 566
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 GTGAACACCAACACATCCTACAAAAGCCTTCAGCGAGCGAGCGGCGGTTACAGCCCCGGTTCGGT 300

Qy          567 CTGCTCGTTCGGTGGGGCCTTGGACGGGCATTAATATTACACTCACAGGGGACCCCAGTGCATCAGCTG 626
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 CTGCTCGTTCGGTGGGGCCTTGGACGGGCATTAATATTACACTCACAGGGGACCCCAGTGCATCAGCTG 360

Qy          627 AACGAGACCAATTGACTACTACAACGAGACGAGTTCACCTGGGCGTTCTGAAAGAGAAATTACGCGCCG 686
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 AACGAGACCAATTGACTACTACAACGAGACGAGTTCACCTGGGCGTTCTGAAAGAGAAATTACGCGCG 420

Qy          687 GAGTACGCGGAACCGCACTGGAGAAGGGGGTGCCGGACCCAGTGCTTCTACCTGGCGGAGAAG 746
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 GAGTACGCGGAACCGCACTGGAGAAGGGGGTGCCGGACCCAGTGCTTCTACCTGGCGGAGAAG 480

Qy          747 TTCACACCCGAGTAGCCCTTTGCGGCTGTACCACCTGGCGCGGGACACTACGGCC 806
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 TTCACACCCGAGTAGCCCTTTGCGGCTGTACCACCTGGCGCGGGACACTACGGCC 540

Qy          807 TCGGGCCACGGCTATGGGTGGGCGTTTCGTTCGTTCGGTCTTCCTCCAAGGTGCTGTGCTCTCCACG 866
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 TCGGCCACGGCTATCGGGTGACGTTTCGTTCGGTCTTCCTCCAAGGTGCTGTGCTCTCCACG 600

Qy          867 CCCGGCCCCGCTCTACGGAGGCCACTGCAGCGAGCCTTCGGCGCTCTTCGGG 926
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 CCCGGCCCCGCTCTACGGAGGCCACTGCAGCGAGCCTTCGGCGCTCTTCGGG 660

Qy          927 GTCTTTCGGCCTTGGCCTTCCATTCTCAGCGGTGCGCGCTCGCCGGCCTCAGGGCTCCC 986
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 GTCTTTCGGCCTTGGCCTTCCATTCTCAGCGGTGCGCGCTCGCCGGCCTCAGGGCTCCC 720

Qy          987 GCGGTCACCACTCAGTACGGCGCGCCCGCGCCTTCGGCCCAAACCGGGCGTCGTGTGC 1046
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          721 GCGCTCACCACTCAGTACGGCGCGCCCGCGCCTTCGGCCCAAACCGGGCGTCGTGTGC 780

Qy         1047 CTCTTCCTCGGAGGGGCCGTGGGTGAGTCTCCAGTATGTTCGGCGCGGCGTCTTCGGCACC 1106
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          781 CTCTTCCTCGGAGGGGCCGTGGGTGAGTCTCCAGTATGTTCGGCGCGGCGTCTTCGGCACC 840

Qy         1107 CTTCTCGGACCAAAGGCGCCAAGGAGCTCAGCGCAGGAGAGGGGGCGTCACCTCTTTATCCTC 1166
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          841 CTTCTCGGACCAAAGGCGCCAAGGAGCTCAGCGCAGGAGAGGGGGCGTCACCTCTTTATCCTC 900

Qy         1167 GGGGACCCACTGCACAAGCAGGGCCCGCCGCCTCTTCCCAGACTTAAAATGTATCACCAACTG 1226
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          901 GGGGACCCACTGCACAAGCAGGGCCCGCCGCCTCTTCCCAGACTTAAAATGTATCACCAACTG 960
```

```
RESULT 11
ACF58262
ID   ACF58262 standard; DNA, 981 BP.
XX
AC   ACF58262;
XX
DT   12-FEB-2004 (first entry)
XX
DE   Human chr15.41.013.a nucleotide sequence.
XX
KW   Colon cancer; cytostatic; gene therapy; colorectal tumour; human; gene;
KW   ds.
```

```
XX
OS   Homo sapiens.
XX
FH   Key              Location/Qualifiers
FT   CDS              1..981
FT                    /*tag= a
XX
PN   WO2003083074-A2.
XX
PD   09-OCT-2003.
XX
PF   28-MAR-2003; 2003WO-US009534.
XX
PR   28-MAR-2002; 2002US-0367727P.
PR   20-MAY-2002; 2002US-0381328P.
PR   10-JUN-2002; 2002US-0386747P.
PR   20-NOV-2002; 2002US-0427564P.
XX
PA   (IDEC-) IDEC PHARM CORP.
XX
PI   Maclachlan K, Gately D;
XX
DR   WPI; 2003-804049/75.
DR   P-PSDB; ABR83679.
XX
PS   Claim 4; Page 92-93; 0pp; English.
XX
PT   New isolated nucleic acid that is expressed by human colon cancer cells,
PT   useful for treating colon cancer, for expressing the corresponding
PT   antigen, and for producing ligands that specifically bind such antigen.
XX
CC   The invention relates to nucleic acids and proteins that are over
CC   expressed in colon cancer cells. The nucleic acids, antigens, primers and
CC   methods are useful for detecting colon cancer. The method and kit provided
CC   are useful for detecting colon cancer. The genes are useful for
CC   expressing the corresponding antigen, and for producing ligands that
CC   specifically bind such antigen, e.g. monoclonal antibodies and small
CC   molecules. The present sequence represents a chr15.41.013.a nucleotide
CC   sequence, overexpressed in colon tumour tissue
XX
SQ   Sequence 981 BP; 171 A; 315 C; 294 G; 201 T; 0 U; 0 Other;

     Query Match             54.1%;  Score 768.8; DB 10;  Length 981;
     Best Local Similarity   99.7%;  Pred. No. 4.9e-182;
     Matches   770;  Conservative    0;  Mismatches    2;  Indels   0;  Gaps    0;

Qy          267 ATGACCCTGTGGAACGGCGTACTGCCTTTTTTACCCCCAGCCCGCATGCGCAGGCTTC 326
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 ATGACCCTGTGGAACGGCGTACTGCCTTTTTTACCCCCAGCCCGCATGCGCAGGCTTC 60

Qy          327 AGCCGTTCCACTGCTCATCGTTATTCTAGTGTTTTTTGGCTCTAGCAGCAAGCTTCCTGCTC 386
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 AGCCGTTCCACTGCTCATCGTTATTCTAGTGTTTTTTGGCTCTAGCAGCAAGCTTCCTGCTC 120

Qy          387 ATCTTGCCGGGGGATCCGGTGGCCACTCGCGCGCGGTGGTTTTGGTTGGTGAGAGTTCTTCTTCAGT 446
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 ATCTTGCCGGGGGATCCGGTGGCCACTCGCGCGCGGTGGTTTTGGTTGGTGAGAGTTCTTCTTCAGT 180

Qy          447 CTGTTCATACGGCGCAGAGAAATGTGGCTGTGTCCACTCCAGTCAGAATGGTTCGTGGGTACA 506
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 CTGTTCATACGGCGCAGAGAAATGTGGCTGTGTCCACTCCAGTCAGAATGGTTCGTGGGTACA 240

Qy          507 GTGAACACCAACACCATCCTACAAAAGCCTTCAGCGCCGCAGCGCGGGCCGTTACAGCCCGTGTCGGT 566
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 GTGAACACCAACACCATCCTACAAAAGCCTTCAGCGCCGCAGCGCGGGCCGTTACAGCCCGTGTCGGT 300

Qy          567 CTGCTCGTGGGCCTGGAGGGCATTAATATTATTACACTTCACAGGGACACCCAGTTGCATCAGCTG 626
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 CTGCTCGTGGGCCTGGAGGGCATTAATATTATTACACTTCACAGGGACACCCAGTTGCATCAGCTG 360

Qy          627 AACGAGACCATTGACTACACAACGAGCAGTTCACCTGGCGTCTGAAAGAGAATTACGCGCGCG 686
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 AACGAGACCATTGACTACACAACGAGCAGTTCACCTGGCGTCTGAAAGAGAATTACGCGCGCG 420

Qy          687 GAGTACGCGGAACGCGCACTGGAGAAGAGGGGGCTGGCCGGACCCAGTGCCTCTACCTGGGCGGAGAAG 746
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 GAGTACGCGGAACGCGCACTGGAGAAGAGGGGGCTGGCCGGACCCAGTGCCTCTACCTGGGCGGAGAAG 480

Qy          747 TTCACACCGAGTAGCCCTTCGCGGCCTGTACCACCAGTACCACCTGGCGCGGGACCAGTACGGCC 806
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 TTCACACCGAGTAGCCCTTCGCGGCCTGTACCACCAGTACCACCTGGCGCGGGACCAGTACGGCC 540

Qy          807 TCGGCCCACCCTATGGGTGCGGTTTGCTTGTCTGGCTCTCCTCTCCCACGTGTCTCTCCACG 866
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 TCGGCCCACCCTATGGGTGCGGTTTGCTTGTCTGGCTCTCCTCTCCCACGTGTCTCTCCACG 600

Qy          867 CCGGCCCCGGCTCTACGGGAGCGGCCTGGCCACTGCCACGCCGGAGGCCTTTCGGCGGCTCTTCGGG 926
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 CCGGCCCCGGCTCTACGGGAGCGGCCTGGCCACTGCCACGCCGGAGGCCTTTCGGCGGCTCTTCGGG 660

Qy          927 GTCTTCGGCCCTTGGCCTTCCATCTCTACGCGGTGGGCGTCTCGCCCGCCCGGCCTCGGGGCTCCTCC 986
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 GTCTTCGGCCCTTGGCCTTCCATCTCTACGCGGTGGGCGTCTCGCCCGCCCGGCCTCGGGGCTCCTCC 720

Qy          987 GCGCTCACACACTCACTACGGCGCGCCCCTTCTGGGGTCACGGCTGGGCAACCGGCG 1038
                |||||||||||||||||||||||||||||||||||||||||||||||
Db          721 GCGCTCACACACTCACTGTACGGCGCGCCCGGCCTTCTGGGGTCACGGCTGGGCAACCGGTG 772
```

RESULT 12
AD085846
ID   AD085846 standard; DNA; 981 BP.
XX
AC   AD085846;
XX
DT   26-AUG-2004 (first entry)
XX
DE   Colon cancer associated gene AI821606 related gene, chr15.41.013.a.
XX
KW   overexpressed; human cancer cell; cytostatic; immunostimulant;
KW   gene therapy; cancer; colon; pancreas; breast; ovary; lung;
KW   immune response; ds; gene.
XX
OS   Unidentified.
XX
PN   WO200046342-A2.
XX
PD   03-JUN-2004.
XX
PF   20-NOV-2003; 2003WO-US037206.
XX
PR   20-NOV-2002; 2002US-0427564P.
XX
PA   (BIOG-) BIOGEN IDEC INC.
XX
PI   Maclachlan K, Gately DP;
XX
DR   WPI; 2004-420627/39.
DR   P-PSDB; AD085847.
XX
PT   New ligands, useful in diagnosing and treating cancer, e.g. cancer of the
PT   colon, pancreas, breast, ovary or lung.
XX
PS   Example 6; SEQ ID NO 54; 133pp; English.
XX
CC   The invention relates to a novel isolated nucleic acids and proteins
CC   overexpressed by human cancer cells. The invention further comprises: a
CC   primer mixture comprising primers that result in the specific
CC   amplification of the isolated nucleic acid; an antigen expressed by human
CC   cancer cells comprising an antigen encoded by the isolated nucleic acid,
CC   an antigen having the sequence of 273 or 234 amino acids, or a fragment
CC   or variant of both antigens; a monoclonal antibody or its antigen-binding
CC   fragment that specifically binds to a cancer antigen; a diagnostic kit;
CC   for detecting cancer, comprising the isolated nucleic acid, or a
CC   monoclonal antibody and a detectable label; a method of detecting cancer;
CC   and a method for treating cancer in a subject. The nucleic acids and

CC   proteins, methods and compositions have cytostatic and immunostimulant
CC   activities. The nucleic acids may be used to treat disorders by gene
CC   therapy. The nucleic acids and proteins, methods and compositions are
CC   useful in diagnosing and treating cancer, e.g. cancer of the colon,
CC   pancreas, breast, ovary or lung, and in inducing an immune response. This
CC   polynucleotide represents a DNA sequence relating to the overexpressed
CC   nucleic acids of human cancer cells of the invention.
XX
SQ   Sequence 981 BP; 171 A; 315 C; 294 G; 201 T; 0 U; 0 Other;

Query Match   54.1%;   Score 768.8;   DB 12;   Length 981;
Best Local Similarity   99.7%;   Pred. No. 4.9e-182;
Matches   770;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

```
Qy   267 ATGACCCTGTGGAACGGCGTACTGCCTCGCC...                   326
         |||||||||||||||||||||||||||||||
Db     1 ATGACCCTGTGGAACGGCGTACTGCC...                         60

Qy   327 AGCGTTCCACTGCTCATCGTTATTC...                         386
         ||||||||||||||||||||||||||
Db    61 AGCGTTCCACTGCTCATCGTTATTC...                        120

Qy   387 ATCTTGCCGGGGATCCGTGGCCACT...                         446
         ||||||||||||||||||||||||||
Db   121 ATCTTGCCGGGGATCCGTGGCCACT...                        180

Qy   447 CTGTTCATAGGCGCAGAAATTGTGG...                         506
         ||||||||||||||||||||||||||
Db   181 CTGTTCATAGGCGCAGAAATTGTGG...                        240

Qy   507 GTGAACACCAACACATCCTACAAAGCC...                       566
         |||||||||||||||||||||||||||
Db   241 GTGAACACCAACACATCCTACAAAGCC...                      300

Qy   567 CTGCTCGTGGGCCTGGAGGGGCATTA...                        626
         ||||||||||||||||||||||||||
Db   301 CTGCTCGTGGGCCTGGAGGGGCATTA...                       360

Qy   627 AACGAGACCATTGACTACAACGAGCA...                        686
         ||||||||||||||||||||||||||
Db   361 AACGAGACCATTGACTACAACGAGCA...                       420

Qy   687 GAGTACGCGAACGCACTCGAGAAGGG...                        746
         ||||||||||||||||||||||||||
Db   421 GAGTACGCGAACGCACTCGAGAAGGG...                       480

Qy   747 TTCACACCCAGTAGCCCTTGGCGCCCG...                       806
         |||||||||||||||||||||||||||
Db   481 TTCACACCCAGTAGCCCTTGGCGCCCG...                      540

Qy   807 TCGGCCACGCTATGGGTGGCGTTCTG...                        866
         ||||||||||||||||||||||||||
Db   541 TCGGCCACGCTATGGGTGGCGTTCTG...                       600

Qy   867 CCGGGCCCCGCTCTACGGAGGCCTTG...                        926
         ||||||||||||||||||||||||||
Db   601 CCGGGCCCCGCTCTACGGAGGCCTTG...                       660

Qy   927 GTCTTGCGCTTGGCCCTTCCATCTCT...                        986
         ||||||||||||||||||||||||||
Db   661 GTCTTGCGCTTGGCCCTTCCATCTCT...                       720

Qy   987 GCGCTCACCACTCAGTACGGGCGCGCC...                      1038
         |||||||||||||||||||||||||||
Db   721 GCGCTCACCACTCAGTACGGGCGCGCC...                      772
```

RESULT 13
AAA94624
ID   AAA94624 standard; DNA; 498 BP.
XX
AC   AAA94624;
XX

DT   11-JAN-2001 (first entry)
XX
DE   Human CASB618 EST.
XX
KW   Epitope; human; CASB618; chromosome 15; tumour; vaccine; cancer; ovarian;
KW   colon; autoimmune disease; HLA_A0201; expressed sequence tag; EST; SB.
XX
OS   Homo sapiens.
XX
PN   WO200053748-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-EP002048.
XX
PR   11-MAR-1999;   99GB-00005607.
PR   01-SEP-1999;   99GB-00020590.
XX
PA   (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Bruck CEM, Cassart J, Coche T, Vinals Y De Bassolac;
XX
DR   WPI; 2000-572268/53.
XX
PT   New human CASB618 polypeptide, useful as a vaccine for prophylactic and
PT   therapeutic treatment of cancers, particularly ovarian or colon cancer,
PT   autoimmune diseases and related conditions.
XX
PS   Claim 32; Page 62; 76pp; English.
XX
CC   The present sequence is an expressed sequence tag (EST) for human CASB618
CC   protein. The gene for human CASB618 protein is thought to be located on
CC   chromosome 15. CASB618 protein and epitopes of CASB618 protein (see
CC   AAB261327 to AAB263399) are useful in diagnosing the occurence of tumour
CC   cells and in vaccines for prophylactic and therapeutic treatment of
CC   cancers, particularly ovarian or colon cancer, autoimmune diseases and
CC   related conditions.
XX
SQ   Sequence 498 BP; 97 A; 166 C; 137 G; 98 T; 0 U; 0 Other;

Query Match   31.9%;   Score 453;   DB 3;   Length 498;
Best Local Similarity   99.6%;   Pred. No. 4e-103;
Matches   475;   Conservative   0;   Mismatches   0;   Indels   2;   Gaps   2;

```
Qy   947 CTCTAGCGTGCCGCCTCTGCCCGGCT-CCGCCTAGGCTCCTCCGCGTCACTCAGTACG   1005
         ||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db     1 CTCTAGCGTGCCGCCTCTGCCCGGCTCCGCCTAGGCTCCTCCGCGTCACTCAGTACG      60

Qy  1006 -GCGCGCCGCCTTCTGGGTGCACCGTGGCAACCGGCGTGCCTGGCCGTTCCTTCG       1064
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AGCGCGCCGCCTTCTGGGTGCACCGTGGCAACCGGCGTGCCTGGCCGTTCCTTCG      120

Qy  1065 GTGGTGAGTCCCAGTATAGTTCGGCGCCAGCGCCTTCGGCACGCTTGGACCAAAGCGCC   1124
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGGTGAGTCCCAGTATAGTTCGGCGCCAGCGCCTTCGGCACGCTTGGACCAAAGCGCC   180

Qy  1125 AAGGACTGCAGCGCCAGGAGAGGGGCTCACCTCTTATCCTCGGGGAGGGGGC         1184
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAGGACTGCAGCGCCAGGAGAGGGGCTCACCTCTTATCCTCGGGGAGGGGGC        240

Qy  1185 CAGGGCGCTCTCCCAGACTTAAAATGTATCACCACTAACCTGTGAGGGGGACCCAAATCTG  1244
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGGGCGCTCTCCCAGACTTAAAATGTATCACCACTAACCTGTGAGGGGGACCCAAATCTG   300

Qy  1245 GACTCCTTCCCCGCCTTGGCACATCGCAGCGCGGGAAAGCAGTGCCCGCCGCCAGGGCGCCC  1304
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GACTCCTTCCCCGCCTTGGCACATCGCAGCGCGGGAAAGCAGTGCCCGCCGCCAGGGCGCCC   360

Qy  1305 AGGAGGAGCTCCAGGAAGGCCACATCGACATCGAGCGGCTGCGGCGGAGGCCTTCGGACATCC  1364
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AGGAGGAGCTCCAGGAAGGCCACATCGACATCGAGCGGCTGCGGCGGAGGCCTTCGGACATCC   420

Qy  1365 ACCAGGGAAAGTCTCCTGGGGGATCTGTAAATAAACCTTTTTTTTTCTTTTGTTTTTT    1421
```

Db 421 ACCAGGGAAAGTCTCCTGGGGCGATCGTAAATAAACCTTTTTTTTCTTTTGTTTTTT 477
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy ...

RESULT 14
AC AB274587
XX
ID AB274587 standard; DNA; 5033 BP.
XX
DT 12-MAY-2003 (first entry)
XX
AC AB274587;
XX
DE Secreted protein gene 360 genomic fragment HUFCL31, SEQ ID NO:1734.
XX
KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;
KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
KW drug screening; chromosome identification; chromosome mapping;
KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
KW antianaemic; vulnerary; chromosome 9p21; gene; ds.
XX
OS Homo sapiens.
XX
PN WO20277013-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-US009370.
XX
PR 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-0095 0082.
PR 12-SEP-2001; 2001US-0095 0083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-040578/03.
XX
PT New human secreted proteins and nucleic acids, useful for detecting or
PT treating cancer or other hyperproliferative disorders, autoimmune
PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
XX
PS Disclosure; Page 2315-2316; 2474pp; English.
XX
CC AB273281-AB273697 represent cDNAs corresponding to 391 human secreted
CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
CC AB273698-AB274687 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins are thought to be involved in biological activities
CC associated with cellular signalling, cellular differentiation, cell
CC migration, prohormone activation and neurotransmitter activity. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing or treating cancers or other
CC hyperproliferative disorders. Additionally, the secreted proteins and
CC their nucleic acids may also be used in the treatment of autoimmune
CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
CC wound healing. Nucleic acids of the invention may be used for chromosome
CC identification, chromosome mapping, in gene therapy, for identifying
CC individuals from minute biological samples, as hybridisation probes, and
CC as molecular weight markers. The present sequence represents a human
CC secreted protein genomic fragment referred to in the disclosure of the
CC invention.
XX
SQ Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;

Query Match   29.1%; Score 413.4; DB 8; Length 5033;
Best Local Similarity 98.6%; Pred. No. 6.7e-93;
Matches 417; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1   GAACCAGGAAAAGTAAACGGCTACAGACACGTGAGAGAAATAGTTTCGTTCCTTAAACGGA 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db 5   GAACCAGGAAAAGTAAACGGCTACAGACAGTGAGAGAAATAGTTTCGCCGGCTTAACGGA 64

Qy 61 CTCTGTCGGTACCAACCCCAGAGACGGTTGAGAGCAGCAGCCCACCTCACAGCAAAAGACG 120
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 65 CTCTGTCGGTACCAACCCCAGAGACGGTTGAGAGCAGCACCTCACAGCTTCCTTAACGGA 124

Qy 121 GAGGTGCAGGACTTCACCAGACCCACTCGTTCCAGCCCTTGTACGGAAAAGACG 180
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125 GAGGTGCAGGACTTCACCAGACCCACTCGTTCCAGCCCTTGTACGGAAAAGACG 184

Qy 181 CCAAGGACGCGGCCCTTCCCCGGTCCAGGTCCAGCCCCAGCAGCCTTGGCTTGGCTGG 240
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 185 CCAAGGACGCGGCTCTCCCCGGTCCAGGTCCAGCGCCCAGCTTGCTTGGCTGGCTTG 244

Qy 241 CGTGCAGCACTCGGCCGCGGCGTGCAGCAGCATGACCCTGTGGAACGGCGTACTGCCTTTTACC 300
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 245 CGTGCAGCACTCGGCCGCGGCGTGCAGCAGCATGACCCTGTGGAACGGCGTACTGCCTTTTACC 304

Qy 301 CCCAGCCCCGGGCATGCCCGGAGGGCTTCAGGCGTTCAGCACTGCTCATCGTTATTCTAGTGTTT 360
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 305 CCCAGCCCCGGGCATGCCCGGAGGGCTTCAGGCGTTCCACTGCTCATCGTTATTCTAGTGTTT 364

Qy 361 TGGGCTCTAGCAGCAAGCTTCCCTGTCATCTTGCCGGGGGATCCGGTGGCCACTCGCGCCTGGT 420
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 365 TGGGCTCTAGCAGCAAGCTTCCCTGTCATCTTGCCGGGGGATCCGGTGGGGTAAGGG 424

Qy 421 TTT 423
       | |
Db 425 TGT 427

RESULT 15
AC ADC21005
XX
ID ADC21005 standard; DNA; 5033 BP.
XX
AC ADC21005;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human secreted protein-related DNA sequence #423.
XX
KW gene therapy; human; secreted protein; haemopoietic disorder;
KW haematological disorder; anaemia; haemophilia; inflammatory disorder;
KW inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
KW leukaemia; wound healing; epithelial cell proliferation disorder;
KW immune disorder; autoimmune disorder; asthmatic disorder;
KW cardiovascular disorder; atherosclerosis; myocarditis;
KW infectious disease; HIV; AIDS; endocrine disorder; diabetes;
KW gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.
XX
OS Homo sapiens.
XX
PN WO20292787-A2.
XX
PD 21-NOV-2002.
XX
PF 26-MAR-2002; 2002WO-US009257.
XX
PR 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-0095 0082.
PR 12-SEP-2001; 2001US-0095 0083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-129287/12.
XX
PT New human secreted proteins and nucleic acid molecules, useful for
PT preparing a diagnostic or pharmaceutical composition for diagnosing,

PT   preventing or treating hematopoietic or hematologic disorders, e.g.
PT   anemia or hemophilia.
XX
PS   Disclosure; SEQ ID NO 959; 1512pp; English.
XX
CC   The invention comprises the amino acid and coding sequences of human
CC   secreted proteins. The DNA and protein sequences of the invention are
CC   useful for detecting, preventing, diagnosing, prognosticating, treating
CC   or ameliorating: haematopoietic or haematological disorders (e.g. anaemia
CC   and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease
CC   and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia);
CC   wound healing and disorders of epithelial cell proliferation; immune
CC   disorders (e.g. autoimmune disorders and asthmatic disorders);
CC   cardiovascular disorders (e.g. atherosclerosis and myocarditis);
CC   infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
CC   and gastrointestinal disorders (e.g. duodenal ulcers and
CC   gastroenteritis). The present DNA sequence was used in the
CC   exemplification of the invention.
XX
SQ   Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;

Query Match   29.1%; Score 413.4; DB 10; Length 5033;
Best Local Similarity   98.6%;
Matches 417; Conservative   0; Mismatches   6; Indels   0; Gaps   0;
Pred. No. 6.7e-93;

Qy   1 GAACCAGGAAAGTAAACGGCTACACAGACAGCAGTGAGAAAATAGTTTCGCTCGCGGGCTTAGAAAAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5 GAACCAGGAAAGTAAACGGCTACACAGACAGCAGTGAGAAAATAGTTTCGCTCGCGGGCTTAGAAAAA 64

Qy   61 CTCTGTCGGTACCAACCCCAGAGGCGGTTGAGAGCAGCAGCCCACTCCAGCGTTCCTTAACGGA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65 CTCTGTCGGTACCAACCCCAGAGGCGGTTGAGAGCAGCAGCCCACTCCAGCGTTCCTTAACGGA 124

Qy   121 GAGGTGCAGGAGACTCAGACTTCACCAGCCCACTCGGTCCCAGCCCTTGTACGCAAAGAGACG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   125 GAGGTGCAGGAGACTCAGACTTCACCAGCCCACTCGGTCCCAGCCCTTGTACGCAAAGAGACG 184

Qy   181 CCAAGGACGCGCTCTCCCGGGCGTTCAGGCAGCCCCACTCGGTCCCAGCCTTGCCTTGCCCGCCTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   185 CCAAGGACGCGCTTCTCCCGGTCCAGGCAGCCCCCAGCTTGCTGGCCTTGCCCGCCTG 244

Qy   241 CGTGCAGCAGCACTCCGCCCGCCGGCGTGCAGCCATGACCCTGTGGAACGGCGTACTGCCCTTTTTACC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   245 CGTGCAGCAGCACTCCGCCCGGCGTGCAGCAATGACCCTGTGAACGGCGTACTGCCCTTTTTACC 304

Qy   301 CCCAGCCCCGGCCATGCCCAGGGCTTCAGGGTTCCACTGCTCATCGGTTATTCTPAGTGTTTT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   305 CCCAGCCCCGGCCATGCCCAGGGCTTCAGGGTTCCACTGCTCATCGTTATTCTPAGTGTTTT 364

Qy   361 TGGCTCTAGCAGCAAGCTTCCTGCTCATCTTGCCGGGGATCCGTGGGCCACTCGCGCTGGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   365 TGGCTCTAGCAGCAAGCTTCCTGCTCATCTTGCCGGGGATCCGTGGGCCACTCGCGTAAGGG 424

Qy   421 TTT 423
          | |
Db   425 TGT 427

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 21:54:27 ; Search time 6251 Seconds
(without alignments)
10635.809 Million cell updates/sec

Title: US-09-937-059-57
Perfect score: 1421
Sequence: 1 gaaccaggaagtaacggct.........cttttttccttgttttt 1421

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	963	67.8	963	11	DQ050095	DQ050095 Homo sapi
	2	701.6	49.4	769	11	DQ050096	DQ050096 Pan trogl
	3	693	48.8	1594	4	AK018569	AK018569 Mus muscu
c	4	657	46.2	1309	4	AK008816	AK008816 Mus muscu
c	5	615.8	43.3	619	8	DN831248	DN831248 MGC1.5.1
c	6	610.6	43.0	617	8	DN831247	DN831247 MGC1.5.1.
c	7	602	42.4	602	8	DN930503	DN930503 MGC1.5.1.
c	8	600.8	42.3	622	8	DN930502	DN930502 MGC1.5.1.
c	9	529.6	37.3	537	7	AI821606	AI821606 nk08all.x
c	10	526.4	37.0	717	7	CK971974	CK971974 4087846 B
c	11	522.2	36.7	703	7	CK979918	CK979918 411631 B
c	12	512	36.0	515	1	AI826629	AI826629 wk35e04.x
c	13	510	35.9	510	1	AI304327	AI304327 qo57h06.x
c	14	496	34.9	500	3	BP396416	BP396416 BP396416
c	15	489.2	34.4	505	1	AI791844	AI791844 nk08all.y
c	16	468.8	33.0	704	7	CK963936	CK963936 4078848 B
c	17	468.2	32.9	886	7	CJ022684	CJ022684 CJ022684
c	18	467.4	32.9	486	5	AI660560	AI660560 we68b08.x
c	19	466	32.8	1014	5	BY708714	BY708714 BY708714
c	20	456.8	32.1	716	7	CK981626	CK981626 CK981626
c	21	450	31.7	849	1	CJ023383	CJ023383 CJ023383
c	22	445.4	31.3	466	1	AA573825	AA573825 nk08all.s
c	23	443	31.2	446	1	AI984141	AI984141 wu21c02.x
c	24	433	30.5	434	1	AI991272	AI991272 wu41h04.x
c	25	432.2	30.4	826	3	BP435528	BP435528 BP435528
c	26	429	30.2	595	8	DN350567	DN350567 LIB3579-0
c	27	418	29.4	876	1	BP440241	BP440241 BP440241
c	28	414	29.1	414	1	AA593860	AA593860 nn19f03.s
c	29	405	28.5	405	1	AI274929	AI274929 ql49c11.x
c	30	404.4	28.5	411	1	AI983793	AI983793 wu20c09.x
c	31	400.2	28.2	405	1	AI346155	AI346155 qp43f12.x
c	32	391	27.5	538	1	AI660493	AI660493 we67h02.x
c	33	389.8	27.4	806	7	CK789912	CK789912 AGENCOURT
c	34	388.8	27.3	538	5	BX517021	BX517021 BX517021
c	35	381	26.8	747	7	CJ027584	CJ027584 CJ027584
c	36	380.4	26.8	573	6	CB854121	CB854121 UI-CP-DUI
c	37	380	26.7	573	6	CB305399	CB305399 UI-CP-EN1
c	38	375.8	26.4	737	6	CF766127	CF766127 CES004870
c	39	373.6	26.3	660	6	BY720747	BY720747 BY720747
c	40	372.6	26.2	646	5	CB171450	CB171450 JID602700
c	41	360.4	25.4	713	6	AY418362	AY418362 Homo sapi
c	42	356.4	25.1	1032	10	CK968084	CK968084 4083252 B
c	43	354.6	25.0	698	7	BP433235	BP433235 BP433235
c	44	352.2	24.8	485	3	CN789601	CN789601 4124079 B
c	45	350.4	24.7	671	7		

ALIGNMENTS

RESULT 1
DQ050095

```
LOCUS       DQ050095                 963 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Homo sapiens HC18097 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ050095
VERSION     DQ050095.1  GI:66903294
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 963)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   (er) PLoS Biol. 3 (6), E170 (2005)
   PUBMED   15869325
REFERENCE   2  (bases 1 to 963)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment. Translation starts at the beginning of
            alignment.
FEATURES             Location/Qualifiers
     source          1..963
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
     gene            <1..>963
                     /locus_tag="HC18097"
ORIGIN
```

Query Match   67.8%; Score 963; DB 11; Length 963;
Best Local Similarity   100.0%;   Pred. No. 6.3e-236;
Matches 963; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
LOCUS       DQ050096                 769 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Pan troglodytes HC18097 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ050096
VERSION     DQ050096.1  GI:66903295
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   PLoS Biol. 3 (6), E170 (2005)
   PUBMED   15869325
REFERENCE   2  (bases 1 to 769)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment. Translation starts at the beginning of
            alignment.
FEATURES             Location/Qualifiers
     source          1..769
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1..>769
                     /locus_tag="HC18097"
ORIGIN
```

```
RESULT 2

DQ050096

Query Match             49.4%;  Score 701.6;  DB 11;  Length 769;
Best Local Similarity   91.5%;
Matches  704;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;
                                   Pred: No. 7.9e-169;
```

```
Qy      267 ATGACCCTGTGGAACGGCGTACTGCCTTTTTACCCCCAGCCCCGGCATGCCGCAGGCTTC 326
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGACCCTGTGGAACGGCGTACTGCCTTTTTACCCCCAGCCCCGGCATGCCGCAGGCTTC 60

Qy      327 AGCGGTTCCACTGCTCATCGTTATTCTAGTGTTTTTGGGCTTCGCAGCAAGCTTCCTGCTC 386
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 AGCGGTTCCACTGCTCATCGTTATTCTAGTGTTTTTGGGCTTCGCAGCAAGCTTCCTGCTC 120

Qy      387 ATCTTGCCGGGGATCCGGTGGCCCACTCGCGCGCCGGTTTTGGTTGGTGAGAGTTCTTCTCAGT 446
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 ATCTTGCCGGGGATCCGGTGGCCCACTCGCGCGCCGGTTTTGGTTGGTGAGAGTTCTTCTCAGT 180

Qy      447 CTGTTCATAGGCGGCCAGAAATTGTGGCTGTGCACTTCACGTGCACAGAATGGTTCGTTCTCTGCTC 506
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 CTGTTCATAGGCGGCCAGAAATTGTGGCTGTGCACTTCACGTGCACAGAATGGTTCGTTCTCTGCTC 240

Qy      507 GTGAACACCAACACATCCTACAAAAGCCTTCAGCGCAGCGGGCCGTTACAGCCCCGTGTCGGT 566
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GTGAACACCAACACATCCTACAAAAGCCTTCAGCGCAGCGGGCCGTTACAGCCCCGTGTCGGT 300

Qy      567 CTGCTCGTGGGCCTGGAGGGCCATTAATATTACACTCACAGGGGACCCCAGTGCATCAGCTG 626
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 CTGCTCGTGGGCCTGGAGGGCCATTAATATTACACTCACAGGGGACCCCAGTGCATCAGCTG 360

Qy      627 AACGAGACCATTGACTACAACGAGCAGTTCACCTGGCGTCTGAAAGAGAATTACGCCGCG 686
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AACGAGACCATTGACTACAACGAGCAGTTCACCTGGCGTCTGAAAGAGAATTACGCCGCG 420

Qy      687 GAGTACGCGGAACGGCACTGGAGAAGGGGCTGCCGGACCCAGTGCTTCTACCTGGCGGAGAAG 746
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GAGTACGCGGAACGGCACTGGAGAAGGGGCTGCCGGACCCAGTGCTTCTACCTGGCGGAGAAG 480

Qy      747 TTCACACCGAGTAGCCCTTGCGCGCCTGTACCACCAGTACCACCTGGCGCGGGACACTACGCC 806
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 TTCACACCGAGTAGCCCTTGCGCGCCTGTACCACCAGTACCACCTGGCGCGGGACACTACGCC 540

Qy      807 TCGGCCACGCTATGGGGTGGCGTTCTGCTTCTCCAAGTGCTGTCTCTCCACG 866
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 TCGGCCACGCTATGGGGTGGCGTTCTGCTTCTCCAAGTGCTGTCTCTCCACG 600

Qy      867 CCGGCCCCGCTCACGCAGGCCTGGACACTGCTACACGGAGAGCCTTGGCGCTCTCGGG 926
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 CCGGCCCCGCTCACGCAGGCCTGGACACTGCTACACGGAGAGCCTTGGCGCTCTCGGG 660

Qy      927 GTCTCTGCCTTGGCCTTCCATCTTCTAGCGGTGCCGCGGTCGCGGTCGCGGGCTCCTCTC 986
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GTCTCTGCCTTGGCCTTCCATCTTCTAGCGGTGCCGCGGTCGCGGTCGCGGGCTCCTCTC 720

Qy      987 GGCGCTCACCACTCAGTACGGCCGTGGTCAGTGTCTGGCCTTCTGGGCTGGCAACCCGGTCCTGTGC 1046
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 GGCGCTCACCACTCAGTACGGCCGTGGTCAGTGTCTGGCCTTCTGGGCTGGCAACCCGGTCCTGTGC 780

Qy     1047 CTCTTCTTCCTCGGAGGGGCCGTGGTGGTCTCCAGTATGTTCGGCCCAGGCGCTCTTCGCACC 1106
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 CTCTTCTTCCTCGGAGGGGCCGTGGTGGTCTCCAGTATGTTCGGCCCAGGCGCTCTTCGCACC 840

Qy     1107 CTTCTGGACCAAAGCGCCAGCAGGACTGCACGAGAGGAGAGGGGCTCACCTTTATCCTC 1166
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 CTTCTGGACCAAAGCGCCAGCAGGACTGCACGAGAGGAGAGGGGCTCACCTTTATCCTC 900

Qy     1167 GGCGACCCACTGCACAAGCAGCGCGCTCTCCCAGACTTAAAGTATACACCACTAACCTG 1226
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 GGCGACCCACTGCACAAGCAGCGCGCTCTCCCAGACTTAAAGTATACACCACTAACCTG 960

Qy     1227 TGA 1229
                |||
Db      961 TGA 963
```

```
RESULT 3
AK018569

LOCUS       AK018569                1594 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus adult male colon cDNA, RIKEN full-length enriched
            library, clone:9030623N16 product:hypothetical protein, full insert
            sequence.
ACCESSION   AK018569
VERSION     AK018569.1  GI:12858338
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The FANTOM Consortium.
  TITLE     The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  JOURNAL   Functional annotation of a full-length mouse cDNA collection
  TITLE     Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1594)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,Y., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.jp/) for further
            details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAAGCAAGGATCCAAGACGTCTTTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 10.0 and subtraction to
            Rot = 105.2. Second strand cDNA was prepared with the primer
            adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..1594
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="9030623N16"
                     /sex="male"
                     /tissue_type="colon"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             144..1106
                     /note="unnamed protein product; hypothetical protein
                     (evidence: ProCrest.decoder,NCBI CDS
                     Predictor,ProCrest.decoder,NCBI CDS
                     putative"
                     /codon_start=1
                     /db_xref="FANTOM_DB:9030623N16"
                     /protein_id="BAB31281.1"
                     /db_xref="GI:12858339"
                     /translation="MTAWDGVLPFYPQPRHAASPSVPLLIVILVFLSLAASPLFLPG
                     IRGHSRWFWLVRVLLSLPIGAEIVAVHFSGDWFVGRVWTNTSYKAFSPSRVQVHVGLH
                     VGIAGVNILRGTPRQQLNETIDYNERFTWRLNEDYTKEVVHALEKGLPDVLIAEK
                     FTPSSPCGLYHQYHLAGHYAAATLWVAFCFWIIIANALLSMPAPLYGGLALLTTGAFTL
                     FGVFAFASISSVPLCHFRLGSAVLFPYGASFWTLATGILSLLLGGAVILHYRPS
                     ALRSFLDLSVRDCSNQAKGNSPLTLNNPQHBEQLKSPDLNTTLL"
ORIGIN
        Query Match        48.8%;  Score 693;  DB 4;  Length 1594;
        Best Local Similarity  75.9%;  Pred. No. 1.4e-166;
        Matches 914;  Conservative   0;  Mismatches 275;  Indels  16;  Gaps   4;

Qy        230 CTGCCCGGCCTTCCGTGCAGCACTCGGCGGCGGCGCTTGCAGCCCTGTGGACGGCGTACT 289
              |||| ||||||||  ||||||||| ||||||| ||||| | ||||||||| || |||||
Db        107 CTCGCTGCCTGGCGAACAGTTCTCGAACTTCTTGCAGGACGACTGCGTGGGGACGGCGTCT 166

Qy        290 GCCTTTTTACCCCCAGCCCCCGGCCATGGCCGCCAGGGCTTCAGCCGTTCCACTGCCTCAT 349
              ||||||||||||||| ||| || ||||||||| ||||||||||||||||||| || | |||
Db        167 ACCCTTTTACCCCCAGCCCCAGCCCCAGGCCATGGCCGCCAGCTTCAGCGTACCGCGTGTTAT 226
```

Middle column sequence alignments:

```
Db    481 TTCACACCGAGTAGCCCTTGCGGCGCTGTACCACCAGTACCACTGGGCGCGGGACACTACGGC
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481 ...

Qy    807 TCCGGCCAGGCTATGGGGTGCGTTCTGCTTCTCGGCCTCTCCAACGTGACTGCTGTCTCCACG 866
          |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db    807 TCCGGCCACCTATGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCACG 866

Qy    541 TCCGGCCACCCTATGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCACG 600
          |||||||||| ||||||||||||||||||||
Db    541 ...

Qy    867 CCGGCCCCGGCCTACGGAGGCCTGGCACTCGGCCGGAGGCCTTCGGCGTCGGCTTTCGGG 926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    867 CCGGCCCCGGCTACGGGAGGCCTGGCACTCGGCCGGAGGCCTTCGGCGTCGGCTTTCGGG 926

Qy    601 CCGGCCCGGCNNNNACGGAGGNNNNAGCACTNNNNNNNNCACCACGGACGGCGTCTTTCGGG 660
          |||||||||||      |||    ||| ||       |||||||| ||| ||  |||||||
Db    601 ...

Qy    927 GTCTTCGCCTTGGGCCTTCCATCTCTAGGCGTCGCCGCTCCGCCTAGGCCTTCGG 986
          ||| ||||| | || |||| |||||||||||||||| ||||||||| ||| |||
Db    927 ...

Db    661 GNNTTNGCCTTGGCCTTCCATCTCCAGGCGTGCGGCGCTCGCCGCNNCGGNNNAGGCTCCTCC 720

Qy    987 GCGGCTCACCACTCAGTACCGGCGCGCGCGCGCGCCTTCTGGGTCACCGCTGGCAACCG 1035
          ||||||||||||||||||||| |||||||||||||||| |||| |||||| ||||| |||
Db    721 GCGGCTCACCACTCAGTATGGCGCGCGCCCCCTNNNGGGTCACCGCTGGCAACCG 769
```

```
Qy   350 TCTAGTGTTTTTGGCTTCTAGCAGCAAGTTCTTGCCGGGGGATCCGTGGCCA 409
         |||| ||||| || |  ||   |   ||  || |||| |   |  ||| ||||
Db   227 CCTGGTGTTCTTTGTCTTTGGCTTCCATCTTGCCTGGAATCCGTGTGGGCA 286

Qy   410 GGCTGCGCACTTCAGTCAGTGCAGAATGGTTCGTTGGGTACAGTGAACACACATCCTACA 469
         |||||||||||||| |||||||||||||| |||||||| ||||||||||||||||||||||
Db   287 CTCGCGCGCTGGTTCTTGGGTTGGTTGGGTAGAGTTCTTCTCAGTGTTT 346

Qy   470 GGCTGTGCACTTCAGTGCACTGGTGAGACTGGGTTCGTTGGGGAGAGTGTGGACCAACACATCCTACA 529
Db   347 GGCTGTGCACTTCAGTGCACTGGTGAGACTGGGTTCGTTGGGGAGAGTGTGGACCAACACATCCTACA 406

Qy   530 AGCCTTCAGCGCAGCGGCCGTTACAGCCCGTTCGTTGGGGAGACTGTGGAGGGCAT 589
         |||||||||||  |||||||||||||||||||||||| |||||||||||||||||
Db   407 AGCCTTCAGCGCCATCCCCTGTTCAAGTCCATGTCGGGTTCGCACGTGTCGTGGGGCGCGT 466

Qy   590 GCAGTTCACCTGGCGTTGAAAGAGAATTACGCGCGCGGAGCTGACTGACTACACAACGA 649
Db   467 TAACATTACACTCCGAGGAACACCGAGGCGCGGAGCTGACTGACTACAACGA 526

Qy   650 GCAGTTCACCTGGCGTTGAAAGAGAATTACGCGCGCGGAGCTGACTGACTACACAACGA 709
Db   527 GCGTTTCACTTGGCGTCTGAACGAGACTACACCAAGGAGATGTCCATGCCTTGGAGAA 586

Qy   710 GGGGCTGCGGGACCCAGTGTGCTTACCTTGGGGAGGAGTTCACCACGGAGTTAGCCCTTGCGG 769
Db   587 GGGACTGCGGGACCCAGTGCTATACCTGGCGAGAGTTCACACCGGAGCTGGTCCTTGCGG 646

Qy   770 CCTGTACCACCAGTACCACCTGGGGGGGACACTAGGCTCCTGGCCGCAACACTCAGTGGTGGCGTT 829
Db   647 GCTGTACCACCAATATCACCTGGCCGGCGAACAAGTCCTACCTGTAGGCTGGGTGGCATT 706

Qy   830 CTGCTTCTGGCTCTCTCTCCAACTGTCTGCTCTCTCCACGCGTCGCCTACCGGAGGGCCT 889
Db   707 CTCGTTCTGGATCATCGGCTGCTCTCTCCCACTCTCCACTGTCTACCGGAGGGCCT 766

Qy   890 GGGCACTGCTGACCACCGGGAGCCTTCGGCCTTCTGGGTTCTTCGCCTTCCATCTC 949
Db   767 GGCTTTGCTCACACACCGGTTCACGCCTTCACGCTCGTCTTCTCTGGCCTTCAAAGA 826

Qy   950 TAGGGTGCCGCTCTGCCTCCGCGCTAGCGTCCGCCTAGGGTCCTCCGGCCGCC 1009
Db   827 CAGGGTGCCGGCTTCTCGCCAGCTTCCGCCACTTCCGCGCTCCTTACTACGGGGC 886

Qy  1010 CGCCTTTCTGGGTCACGGTGGCAACCGGGTCGCCCGCGCTCCTTCCTCGGAGGGCCGTGGT 1069
Db   887 CTCCTTTTTGGCGCTACGGCTGGCCCACGGCGCATCCTGGAGGCCTCCGCTGGGGGTGGT 946

Qy  1070 GAGTCGCTCCAGTATGTTCGGCCCCAGCCGCTCTTCTGGACCAAAAGCGCCCAAGGA 1129
Db   947 GATTCTCCACTATACTCGGCGCCCAGCGCCGTCTTCGCGTTCTCTGGATCTAAGTGTCAAAGA 1006

Qy  1130 CT----GCACCCAGGGAGGAGAGAGGGCACTGAGGCGCTTCACTCCTGGGGGGCCGACCCCAACCGCACACGACGA 1186
Db  1007 CTGTGTAGGCAACCAGGCTAAAAGGGAAACTCACCCTCACCCTCAACAAACCGCAACACGAACA 1066

Qy  1187 GGCGGCGCTCTCCCAGACTTAAAAATGTATCACCACTAACTGTTGACGGGGGGACCCAAATCTGGA 1246
Db  1067 GTTGAAGAGTCCAGACTTAAA---TATTACACTCTCCTGTGAAAATAAACTCTCCGGA 1123

Qy  1247 CTCCCTTCCCCGCGCCCTTGGGACGTGCCGGCCGCAGCCTG-GGCCA 1305
Db  1124 TTTCTACCCCTCTTTGGGACCCCCATAGAACCTGTGTTAAGAGACTGTGCCAGGA 1183

Qy  1306 GGAGAGCTCCAGGAAGGGCACTGAGGCGTCGGCCTCGGACATCCGGCAGGCA 1365
Db  1184 GCATAGCTTGAGTCTGGAGAACTGCTTGTGTCCTGGGGACAGAGGGCGCACCCCATGTG 1243

Qy  1366 CC------AGGGAAAGTCTCCTCTGGGGGGATCTGTAGGGTAAAATAAAAACCTTTTTGTCTTTTGT 1416
Db  1244 CCTTTGCTCTAGAGAATGGGATTTCTCAGGAGGAAGAACTGTAAAATAAAACTTTTTTGTTTTCTTTT 1303

Qy  1417 TTTTT 1421
         |||||
Db  1304 TTTTT 1308
```

RESULT 4

Db AK008816   1309 bp   mRNA   linear   HTC 03-APR-2004
LOCUS   AK008816
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched
            library, clone:2210403N03 product:hypothetical protein, full insert
            sequence.
ACCESSION   AK008816
VERSION   AK008816.1 GI:12843238
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.

REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 1309)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,Y., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,K.,
            Tejima,K., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  JOURNAL   Direct Submission
  TITLE   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,

COMMENT   URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
           Fax:81-45-503-9216)
           Please visit our web site (http://genome.gsc.riken.jp/) for further
           details.
           cDNA library was prepared and sequenced in Mouse Genome
           Encyclopedia Project of Genome Exploration Research Group in Riken
           Genomic Sciences Center and Genome Science Laboratory in Riken.
           Division of Experimental Animal Research in Riken contributed to
           prepare mouse tissues. First strand cDNA was primed with a primer
           [5', GAGAGAGAGAGAGACTCCAAGAGCTCTTTTTTTTTTTTTTTTTTTTVN 3'], cDNA was
           prepared by using trehalose thermo-activated reverse transcriptase
           and subsequently enriched for full-length by cap-trapper. Second
           strand cDNA was prepared with the primer adapter of sequence [5'
           GAGAGAGATTCTCGAGTTAANTTAANTTAANTTCCCCGCCCCCCCCC 3'], cDNA was cleaved
           with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
           Host: SOLR.

FEATURES   Location/Qualifiers
source   1..1309
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:2210403N03"
                     /db_xref="taxon:10090"
                     /clone="2210403N03"
                     /sex="male"
                     /tissue_type="stomach"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
CDS   142..1104
                     /note="unnamed protein product; hypothetical protein
                     (evidence: ProCrest, decoder, NCBI CDS
                     Predictor, Longest-ORF)
                     putative"
                     /codon_start=1
                     /protein_id="BAB25910.1"
                     /db_xref="GI:12843239"
                     /translation="MTAWDGVLPFYPQPRHAASFSVPLLIVIVFLSLAASPLFILPG
                     IRGHSRWFWLVRVLLSLPFIGAEIVAVHFSGDWWTVEEBVWTNTSYXAFSPSRVQVHVGLH
                     VGLAGVNTTLRGTPRQQLNETIDYNERPTWRLNEDYTKEYVHAFEKGLPDPVLYLAEK
                     FTPSSPCGLYHQVHLPGHYAAATLMVGILLLDHRQCAANLHARPTLRRLGFAHHRCLRA
                     LRCLRFRLDFQDRAALPLPPWLRRLHALLRRLLILTTLATGILSLLLGGAVVILHYTRPS
                     ALRSFLDLSVKDCSNQAKGNSPLTLNNPQHEQLKSPDLNITTLL"

ORIGIN

Query Match   46.2%;   Score 657;   DB 4;   Length 1309;
Best Local Similarity   74.5%;   Pred. No. 2.4e-157;
Matches   898;   Conservative   0;   Mismatches 290;   Indels   17;   Gaps   5;

Qy   230 CTGCCCGCCTGGTTTTGGTTGGTGGAGAGTTCTTCTCAGTCTGTCGTTCATTAGGCGCGGAAATTGT 469
           ||||||||||||||||||||||||||||||||||||||||||||| ||||||| |||
Db   465 TAACATTACACTCCGAGGAACACCGAGCGCAGCTGAACGCAGCGCAGATTGACTACAATGA 524

Db   285 CTCGGCGCTGGTTCTCGGTTGGTTGGAGACTCCTTCTTAGCCTGTTCATTAGGTGCCGGAAATTGT 344
           |||| ||||| ||||||||||||||||||||||||||||||||||||||
Qy   650 GCAGTTCACCTTGGCGTCTGAAAAGACACCGAGCAGCTGAACGCGGTA 709

Qy   410 GGCTGTGCACTTCAGTGCACAGAATGGTTCGTGGGTACAGTGAACCACCAACACTCCTACAA 529
           |||||||| ||||||||||| |||||||||||
Db   525 GCGGTTTCACTTGGCGTCTGAACATGAACTAGCACCACCAAGGAGGTAINGTCCATGCCTTTGAGAA 584

Db   105 CTCGGCTGCCTGGGAACAGTTCTCGAACTTCTTGCAGGATCACGTTGCCTGGCTGCAC 164
           ||||||| || |||| || ||||
Qy   710 GGGGCGTGCCCGACCCCAGTGCTGCTCTCTACCTGGCGGGGAGAAGTTCACACCGAGTTACCCTTGCGG 769

Qy   290 GCCTTTTTACCCCAGCCCCCGGCATGCCGCAGGCTTCAGGTACAGTGTGGACCAACACATCCTACAA 349
           |||||||||||| ||||||||||||||
Db   585 GGGGCGTGCCCGACCCCAGTGCTTATACCTGGCGAGAGAAGTTCACACCGACAGTCCTTGCGGG 644

Db   165 ACCCTTTTACCCCAGCCCCAGGCCATCGCCGGCTACCGCGGCTTCAGGTACATGTTTAT 224
           ||||||||||||||||||||||||||||||||||||||
Qy   770 CCTGTACCACCGAGTACCACCTGGCGGGGACAACTAGGCTCGGCCACGCCTATGGG-TGGCGT 828

Qy   350 TCTAGTGTTTTTTGGCTACAGCAAGCTTCCTGCTCATCTTGCCCGGGGGATCCGTGGCCA 409
           ||| |||||| ||| |||||||||||||||||||
Db   645 GCTGTACCACCAATATCACTTGCCCGGTCACTAATGCCGCAGCAACACTGTGGGTTGGCAT 704

Db   225 CCTCGGTGTTCTTTGGCTGCTAGCTTCCTTCCCTGTTCATCTTTGCCTGGAATCCGTGGGCA 284
           ||||| ||||||||||||||||||||||||||||||||||||||
Qy   829 TCTGCTTCTGGCCTCCTCTTCCCAAACGTGTGCTGCTTCTCCAAGCTGCTCTTACGGAGGCT 888

Qy   470 GGCTGTGCACTTCAGTGCACAGAATGGTTCGTGGGTACAGTGAACCACCAACACTCCTACAA 529
           |||||||| ||||||||| |||||||||||||||
Db   765 TGGGTTTTGCTCACCACCGGGTCCCTTCGGCGTCTTGCGGTCTTGCGGTCTTTGGCTTCGATTT 824

Db   345 GGCTGTGCACTTCAGTGCACAGACTGGTTCGTGGGACCAACACATCCTACAA 404
           |||||||||||||||||||||||||||||||||||
Qy   889 TGGCCACTGCTGACCACCGGAGCCTTCGGCGTCTTCGGGTTCGGCCTCGATTT 948

Qy   530 AGCCTTCAGCGCGCCAGCCGCGCCGTTACACCCCGTCTGCTCGTCGGGCCTGGAGGGCAT 589
           |||| |||||||||||||||||||||||||||||||||||
Db   405 AGCCTTCAGTCCATCCCGTTCAAGTCCATGTCGGTCTGCACGTGGGCGTCTGGCGCGCGT 464

Qy   590 TAATATTACTCACAGGAGACCCAGTCGCATCAGCTGAACCGAGACCAATTGACTACAACGA 649
           ||||||||||||||||||||||||||||||||||||||| ||||
Db   465 TAACATTACACTCCGAGGAACACCGAGCGCAGCTGAACGCAGCGCAGATTGACTACAATGA 524

Db   650 GCAGTTCACCTTGGCGTCTGAAAAGACACCGAGCAGCTGAACGCGGTA 709

Db   525 GCGGTTTCACTTGGCGTCTGAACATGAACTAGCACCACCAAGGAGGTAINGTCCATGCCTTTGAGAA 584

Qy   710 GGGGCGTGCCCGACCCCAGTGCTGCTCTCTACCTGGCGGGGAGAAGTTCACACCGAGTTACCCTTGCGG 769

Db   585 GGGGCGTGCCCGACCCCAGTGCTTATACCTGGCGAGAGAAGTTCACACCGACAGTCCTTGCGGG 644

Qy   770 CCTGTACCACCGAGTACCACCTGGCGGGGACAACTAGGCTCGGCCACGCCTATGGG-TGGCGT 828

Db   645 GCTGTACCACCAATATCACTTGCCCGGTCACTAATGCCGCAGCAACACTGTGGGTTGGCAT 704

Qy   829 TCTGCTTCTGGCCTCCTCTTCCCAAACGTGTGCTGCTTCTCCAAGCTGCTCTTACGGAGGCT 888

Db   705 TCTGCTTCTGGATCATCGCCAGCTGCTCGTCTCCCATGCGTGCGCCCGCGCCCGCTCTACGGAGGCT 764

Qy   889 TGGCCACTGCTGACCACCGGAGCCTTCGGCGTCTTCGGGTTCGGCCTCGATTT 948
           || |||||||||| |||||||||| ||||| |||||| |||||||| ||||||
Db   765 TGGGTTTTGCTCACCACCGGGTCCCTTCGGCGTCTTGCGGTCTTTGGCTTCGATTT 824

Qy   949 CTAGCGGTGCCGCTCTTCGCCCGGCTCCGCCTAGGCTCTCCACCACTCAGTACGGCG 1008
           |||||||||||||||||| |||| ||||||||||||||||||||||||||||||||
Db   825 CCAGCGGTGCCGCCTCTTCGCCCACTCCCGGCTCCGCCTTGGCTCCGCCCGGTCTTCACGCCTTACTACGGCG 884

Qy   1009 CCGCCTTCTGGGGTCACGCTGGCCAACCGGCGGTTCCTGTGCCCGCCCCGCCCGGTCGGGTGG 1068
           |||||||||||||||||||| |||||||||||||||||||||||
Db   885 CCTCCCTT-TTGCTCAGCCTGGCCACCGGGCATCCTGAACCGTCCGCCGGTTGGCTGCGGTGG 943

Qy   1069 TGAGTGTCTCCAGTATGTTCGACGGGCTCAGCCCAGGCGCTCCTTATCCTCGCGGGACCCACTGCACAAGC 1128
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   944 TGATTCTCCCACTATACTCGGCCGCGCCCCAGGCCGTGCCCCTTCGGCCTCCTTTCTGGATCCTAAGTGTCAAAG 1003

Qy   1129 ACT----GCAGCCAGGAGAGAGGGGGGCTCACCTCTTTATCCTCGGGGCCGACCACTGCACAAGC 1185
           ||| || |||||||||||||||
Db   1004 ACTGTAGCAACCAGGCTAAAGGGAAACTCACCTCTCACCCTCAACAACCCCGAACACACGAAC 1063

Qy   1186 AGGGCCGCTCTCCCAGACTTAAAATGTATCACCACTAACCTTCTGAGGGGGGACCCAAATCTGG 1245
           |||||||||| |||||||||||||||||||||||||||||
Db   1064 AGTTGAAGAGTCCAGACTTAAA---TATTACCACTCTCCTGTGAAAATAACTGACTCCGG 1120

Qy   1246 ACTCCTTCCCCGCCCTTGGGACATCGGACGGGGGAAGCAGTGCCGCGGGGGGCACAAGC 1305
           |||||||||||||||||||||||||||||||||||||||
Db   1121 ATTTCTTCACCCCTCTTTGGGACCCCATAGACCTGAAGACTGTGTTAAGAGCGGGGTGCCAGG 1180

Qy   1306 GGAGAGCTCCAGGAACGGCACTGAAGCGCCTGCTGGGCGGCGGAGGCCCTCGGACATCGCAGGCA 1365
           ||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   1181 AGCATAGCTTAGTCTGGAGACTGTCTTCGTGCGGGGACGACAAGGGGCGCACCCCATCGTG 1240

Qy   1366 CC--------AGGGAAAGTCTCTCGGGGGCGATCTGTGGGGGCATCTAAAATAAAACCTTTTTTTTTGT 1416
           || ||||||||| |||||||||||
Db   1241 CCTTTGCTCTACAGAATGGAATTTCTCAGGAGAGAACTGTAAAATGAACTTTTTTGTTTTCTTTT 1300

Qy   1417 TTTTTT 1421
           |||||
Db   1301 TTTTTT 1305

RESULT 5
LOCUS   DN831248   619 bp   mRNA   linear   EST 27-APR-2005
DEFINITION   DN831248 MGC1.5.1.1.1.A11.R.1 NIH MGC 331 Homo sapiens cDNA clone
             MGC1.5.1.1.1.A11, mRNA sequence.
ACCESSION   DN831248
VERSION   DN831248.1   GI:62639819
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominidae; Homo.

```
REFERENCE    1  (bases 1 to 619)
AUTHORS      Stevens,M., Wei,C., Gross,S.S., McPherson,J. and Brent,M.R.
TITLE        Exhaustive RT-PCR and sequencing of all novel TWINSCAN predictions
             in human
JOURNAL      Unpublished (2005)
COMMENT      Contact: Brent, MR
             Laboratory of Computational Genomics
             Washington University
             Campus Box 1045, One Brookings Dr., St. Louis, MO 63130, USA
             Tel: (314) 935-6621
             Email: brent@cse.wustl.edu
             All products are PCRed from first-strand cDNA made from pooled mRNA
             sources. PCR product is sequenced directly, typically from the PCR
             primers. All submitted sequences produced a high quality, spliced
             alignment to the genome or assembled into a contig that aligned to
             the genome. This alignment was not necessarily to the targeted
             locus. Legitimate, spliced sequences that resulted from mispriming
             to non-targeted cDNAs were also submitted. All data on all
             experiments can be found at http://genes.cse.wustl.edu/rtdb/.
FEATURES             Location/Qualifiers
     source          1..619
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC1.5.1.1.1.All"
                     /tissue_type="pooled"
                     /clone_lib="NIH MGC 331"
                     /note="tissues: adrenal gland, bone marrow, brain
                     cerebellum, brain (whole), fetal brain, fetal liver,
                     heart, kidney, liver, lung (whole), placenta prostate,
                     salivary gland, skeletal muscle, testis, thymus, thyroid
                     gland trachea, uterus, spinal cord. First-strand cDNA
                     made from pooled mRNA sources. PCR product is sequencing
                     directly, typically from the PCR primers."
ORIGIN
```

```
Query Match      43.3%;  Score 615.8;  DB 8;  Length 619;
Best Local Similarity   99.7%;  Pred. No. 8.2e-147;
Matches   617;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy     302  CCAGCCCCGGCAtNGCCCGCGAGCGTTCAGCGGTTCCACTCGTTCATCGTTATTCTAGTTGTTTTT  361
            |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     619  CCAGCCCCGGCAtGCCCGCGAGCGTTCAGCGGTTCCACTCGTTATTCTAGTGTTTTT  560

Qy     362  GGGCTCTAGCACAGCAAGCTTCCTGCTCATCGTTGCGGGGATCCGTGGCCACTCGGCCTGGGTT  421
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     559  GGGCTCTAGCACAGCAAGCTTCCTGCTCATCCTTGCGGGGATCCGTGGCCACTCGGCCTGGGTT  500

Qy     422  TTGGTTGGTGAGAGTTCTTCTTCAGTCTCGTTCATAGGCCCGCCAGAAATTGTTGGCGTGCACTT  481
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     499  TTGGTTGGTGAGAGTTCTTCTTCAGTCTCGTTCATAGGCCCGCCAGAAATTGTTGGCGTGCACTT  440

Qy     482  CAGTGCAGAATGGTTCGTTGGGGTACAGTGAACAACACATCCTACAAAGCCTTCAGCGC  541
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     439  CAGTGCAGAATGGTTCGTTGGGGTACAGTGAACAACACATCCTACAAAGCCTTCAGCGC  380

Qy     542  AGGGCGCGGTTACAGCCCCGGTGTCGGGTCTGGTCGTTGGAGGGCAATTAATTATTACT  601
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     379  AGGGCGCGGTTACAGCCCCGTGTCCGGTCGTTGGGTGGAGGGCAATTAATTATTACT  320
```

```
RESULT 6
DN831247
LOCUS       DN831247               617 bp    mRNA    linear   EST 27-APR-2005
DEFINITION  MGC1.5.1.1.1.All.F.1 NIH MGC 331 Homo sapiens cDNA clone
            MGC1.5.1.1.1.All, mRNA sequence.
ACCESSION   DN831247
VERSION     DN831247.1  GI:62639816
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 617)
AUTHORS     Stevens,M., Wei,C., Gross,S.S., McPherson,J. and Brent,M.R.
TITLE       Exhaustive RT-PCR and sequencing of all novel TWINSCAN predictions
            in human
JOURNAL     Unpublished (2005)
COMMENT     Contact: Brent, MR
            Laboratory of Computational Genomics
            Washington University
            Campus Box 1045, One Brookings Dr., St. Louis, MO 63130, USA
            Tel: (314) 935-6621
            Email: brent@cse.wustl.edu
            All products are PCRed from first-strand cDNA made from pooled mRNA
            sources. PCR product is sequenced directly, typically from the PCR
            primers. All submitted sequences produced a high quality, spliced
            alignment to the genome or assembled into a contig that aligned to
            the genome. This alignment was not necessarily to the targeted
            locus. Legitimate, spliced sequences that resulted from mispriming
            to non-targeted cDNAs were also submitted. All data on all
            experiments can be found at http://genes.cse.wustl.edu/rtdb/.
FEATURES             Location/Qualifiers
     source          1..617
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC1.5.1.1.1.All"
                     /tissue_type="pooled"
                     /clone_lib="NIH MGC 331"
                     /note="tissues: adrenal gland, bone marrow, brain
                     cerebellum, brain (whole), fetal brain, fetal liver,
                     heart, kidney, liver, lung (whole), placenta prostate,
                     salivary gland, skeletal muscle, testis, thymus, thyroid
                     gland trachea, uterus, spinal cord. First-strand cDNA
                     made from pooled mRNA sources. PCR product is sequencing
                     directly, typically from the PCR primers."
ORIGIN
```

```
Query Match      43.0%;  Score 610.6;  DB 8;  Length 617;
Best Local Similarity   99.4%;  Pred. No. 1.8e-145;
Matches   613;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

Qy     341  CATCGTTATTCTAGTGTTTTTGGCTTGGTGTTGGTTGGTGAGAGTTCTTCTTCAGTCTCGTTCATAGGCGC  400
            |||||||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  CATCGTTATTCTAGTGTTTTTGGCTTCAGCGGCCTTGGTGTTGGTTGGTGAGAGTTCTTCTTCAGTCTCGTTCATAGGCGC  60

Qy     401  CCGTGGGCCACTCGCGGCTGGTTTTTGGTTGGTGGTTGGTGAGAGTTCTTCTTCAGTCTCGTTCATAGGCGC  460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  CCGTGGGCCACTCGGCTGGTTTTTGGTTGGTGGTTGGTGAGAGTTCTTCTTCTTCAGTCATAGGGGC  120
```

```
Db     139  GTACCACCTGGCGGGGGACACTACCGCCTCGGCCACCGCTATGGGGTGGCGGGTTCTGCTTCTGGCT  80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     842  CCTCTCCAACGTGCTGCTCTCCACGCCGCGGCCGCCGCCCCGCTCTACGGGAGCGCCTGGACTGCTGAC  901
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      79  CCTCTCCAACGTGCTGCTCTCCACGCGCGCCGCCGCCCCGCTCTACGGGAGCGCCTGGACTGCTGAC  721

Qy     902  CACCGGAGCCTTCGCGCTC  920
            |||||||||||||||||
Db      19  CACCGGAGCCTTCGCGCTC  1
```

```
Qy     139  GGCTCTAGCACAGCAAGCTTCCTGCTCATCGTTGCGGGGATCCGTGGCCACTCGGCCTGGGTT  200
Db     559  ...
```

RESULT 7
LOCUS   DN930503/c
DEFINITION   MGC1.5.1.2.1.A11.R.1 NIH_MGC_331 Homo sapiens cDNA clone
               MGC1.5.1.2.1.A11, mRNA sequence.
ACCESSION   DN930503
VERSION   DN930503.1 GI:62924189
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Hominidae; Homo.
REFERENCE   1 (bases 1 to 602)
  AUTHORS   Stevens,M., Wei,C., Gross,S.S., McPherson,J. and Brent,M.R.
  TITLE   Exhaustive RT-PCR and sequencing of all novel TWINSCAN predictions
               in human
  JOURNAL   Unpublished (2005)
COMMENT   Contact: Brent, MR
               Laboratory of Computational Genomics
               Washington University
               Campus Box 1045, One Brookings Dr., St. Louis, MO 63130, USA
               Tel: (314) 935-6621
               Email: brent@cse.wustl.edu
               All products are PCRed from first-strand cDNA made from pooled mRNA
               sources. PCR product is sequenced directly, typically from the PCR
               primers. All submitted sequences produced a high quality, spliced
               alignment to the genome or assembled into a contig that aligned to
               the genome. This alignment was not necessarily to the targeted
               locus. Legitimate, spliced sequences that resulted from mispriming
               to non-targeted cDNAs were also submitted. All data on all
               experiments can be found at http://genes.cse.wustl.edu/rtdb/.
FEATURES   Location/Qualifiers
     source   1..602
                     /organism="Homo sapiens"
                     /mol_type="mRNA"

LOCUS   DN930503   602 bp   mRNA   linear   EST 27-APR-2005

ORIGIN

/db xref="taxon:9606"
/clone="MGC1.5.1.2.1.A11"
/tissue_type="pooled"
/clone_lib="NIH_MGC_331"
/note="Tissues: adrenal gland, bone marrow, brain,
cerebellum, brain (whole), fetal brain, fetal liver,
heart, kidney, liver, lung (whole), placenta prostate,
salivary gland, skeletal muscle, testis, thymus, thyroid
gland trachea, uterus, spinal cord. First-strand cDNA
made from pooled mRNA sources. PCR product is sequencing
directly, typically from the PCR primers."

Query Match   42.4%; Score 602; DB 8; Length 602;
Best Local Similarity 100.0%; Pred. No. 2.9e-143;
Matches 602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 8
LOCUS   DN930502
DEFINITION   MGC1.5.1.2.1.A11.F.1 NIH_MGC_331 Homo sapiens cDNA clone
               MGC1.5.1.2.1.A11, mRNA sequence.
ACCESSION   DN930502
VERSION   DN930502.1 GI:62924188
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

LOCUS   DN930502   622 bp   mRNA   linear   EST 27-APR-2005

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE   1 (bases 1 to 622)
AUTHORS   Stevens,M., Wei,C., Gross,S.S., McPherson,J. and Brent,M.R.
TITLE   Exhaustive RT-PCR and sequencing of all novel TWINSCAN predictions
            in human
JOURNAL   Unpublished (2005)
COMMENT   Contact: Brent, MR
            Laboratory of Computational Genomics
            Washington University
            Campus Box 1045, One Brookings Dr., St. Louis, MO 63130, USA
            Tel: (314) 935-6621
            Email: brent@cse.wustl.edu
            All products are PCRed from first-strand cDNA made from pooled mRNA
            sources. PCR product is sequenced directly, typically from the PCR
            primers. All submitted sequences produced a high quality, spliced
            alignment to the genome or assembled into a contig that aligned to
            the genome. This alignment was not necessarily to the targeted
            locus. Legitimate, spliced sequences that resulted from mispriming
            to non-targeted cDNAs were also submitted. All data on all
            experiments can be found at http://genes.cse.wustl.edu/rtdb/.
FEATURES   Location/Qualifiers
     source   1..622
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC1.5.1.2.1.A11"
                     /tissue_type="pooled"
                     /clone_lib="NIH_MGC_331"
                     /note="tissues: adrenal gland, bone marrow, brain
                     cerebellum, brain (whole), fetal brain, fetal liver,
                     heart, kidney, liver, lung (whole), placenta prostate,
                     salivary gland, skeletal muscle, testis, thymus, thyroid
                     gland trachea, uterus, spinal cord. First-stranded cDNA
                     made from pooled mRNA sources. PCR product is sequencing
                     directly, typically from the PCR primers."
ORIGIN

Query Match   42.3%; Score 600.8; DB 8; Length 622;
Best Local Similarity   98.1%; Pred. No. 5.8e-143;
Matches 608; Conservative   0; Mismatches   12; Indels   0; Gaps   0;

Qy   332 TCCACTGCTCATCGTTATTCTAGTGTTTTTTGTGGCTCTAGCGCCAAGCTTCCTGTCATCTT 391
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   2 TTCACTGCTCATCGTTATTCTAGTGTTTTTTGTGGCTCTAGCAGCAAGCTTCCTGTCACCTT 61

Qy   392 GCCGGGGGATCCGTGGCCACTCGCCTGGTTTCGTTGGTGAGGAGTCTTCTTCTCAGTCGTT 451
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   62 TAAGGGCATCCGTGGCCACTCGCCTGGTTTTCGTTAGTTCTTCTCAGTCTCGTT 121

Qy   452 CATAGGGGCCAGAAATTGTGGCGTGTGCACTTCAGTGCAGAAATGTGTTCGTCAGTGAA 511
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 CATAGGGGCCAGAAATTGTGGCGTGTGCACTTCAGTGCAGAATGGTTCGTACAGTGAA 181

Qy   512 CACCAAACACATCCTACAAAAGCCTTCAGCGACGGCAGCGCGGCCCGTCCCGTCGGTCTCGCT 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 CACCAAACACATCCTTACAAAAGCCTTCAGCGACGGCAGCCCGTCCCGTCGGTCTCGCT 241

Qy   572 CGTGGGGCCTGGAGGGCATTAATATATACACTCACAGGGAGCCCAGTGCATCAGCTGAACGA 631
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 CGTGGGGCCTGGAGGGCATTAATTATACACTCACAGGGACCCCAGTGCATCAGCTGAACGA 301

Qy   632 GACCATTGACTACACACGAGCAGTTCACCTGGCGTCTCGAAAAGAAATTACGCCCGCGAGTA 691
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 GACCATTGACTACACACGAGCAGTTCACCTGGCGTCTGAAAGAGAGAAATTACGCCCGGAGTA 361

Qy   692 CCGGAACCGCACTGGAGAAGGGGCTGCCGGACCCAGTGCTCTACCTGCGGAGAAGTTCAC 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 CCGGAACCGCACTGGAGAAGGGGCTGCCGGACCCAGTGCTCTACCTGCGGAGAAGTTCAC 421

Qy   752 ACCGGAGTAGCCCTTGCGGCGGCTGACTACCACCAGTAGCTAGCCTCGGC 811
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422 ACCGGAGTAGCCCTTGCGGCGGCTGACTACCACCAGTAGCTAGCCTCGGC 481

RESULT 9
LOCUS   AI821606   537 bp   mRNA   linear   EST 13-DEC-1999
DEFINITION nk08a11.x5 NCI_CGAP_Co2 Homo sapiens cDNA clone IMAGE:1012892 3'
            similar to contains TAR1.t3 TAR1 MER22 repetitive element i; mRNA
            sequence.
ACCESSION   AI821606
VERSION   AI821606.1 GI:5440685
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 537)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Other ESTs: nk08a11.y5
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
            Ph.D.
            cDNA Library Arraying: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            This read is a RESEQUENCE of a previously sequenced human clone
            Original clone citation: National Cancer Institute, Cancer Genome
            Anatomy Project (CGAP), Tumor Gene Index
            This read has been verified (found to hit its original self in the
            correct orientation)
            Insert Length: 741   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 459.
FEATURES   Location/Qualifiers
     source   1..537
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1012892"
                     /tissue_type="tumor"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="NCI_CGAP_Co2"
                     /note="Organ: colon; Vector: Bluescript SK-; Site 1:
                     EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
                     Oligo dT; Bulk colon villous adenoma. 5' adaptor sequence:
                     5' GAATTCGGCACGAG 3' 3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTT 3' Average insert size: 1.1 kb."
ORIGIN

Query Match   37.3%; Score 529.6; DB 1; Length 537;
Best Local Similarity   99.3%; Pred. No. 1.1e-124;
Matches 532; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   812 CACGCTATGGGTGCGGTTCTTCGCTTCTTGGCTCTTCGGCCTTTCTTCCTTCCAACGTGTGCTCTTCCACGCACCGGC 871
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   482 CACGCTATGGGTGCGGTTCTTCGCTTCTTGGCTCTTCCAACGTGTGCTGCTCTTCCACGCACCGGC 541

Qy   872 CCCGCTCTACGGAGGAGCCTTCGGCACTGCCTGACCACCGGAGCCTTCGCCGTCCTTCGGGGTCTT 931
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   542 CCCGCTCTACGGAGGAGCCTTCGGCACTGCCTGACCACCGGAGCCTTCGCCGTCCTTCGGGGGTCTT 601

Qy   932 CGGCCTTGGCCTCCATTCTTA 951
              |||||||||||||||||||||
Db   602 CGGCCTTGGCCTCCATTCTTA 621

```
RESULT 10
CK971974
LOCUS        CK971974          717 bp          mRNA          linear          EST 16-MAR-2004
DEFINITION   4087846 BARC 9BOV Bos taurus cDNA clone 9BOV4_E23 5', mRNA
             sequence.
ACCESSION    CK971974
VERSION      CK971974.1          GI:45489948
KEYWORDS     EST.
SOURCE       Bos taurus (cow)
ORGANISM     Bos taurus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
             Pecora; Bovidae; Bovinae; Bos.
REFERENCE    1  (bases 1 to 717)
AUTHORS      Sonstegard, T.S., Van Tassell, C.P., Matukumalli, L.K., Harhay,
             G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
TITLE        Production of EST from cDNA libraries derived from immunologically
             activated bovine gut
JOURNAL      Unpublished (2004)
COMMENT      Contact: Tad S. Sonstegard
             Bovine Functional Genomics Laboratory
             Animal and Natural Resources Institute
             Bldg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
             Tel: 301504816
             Fax: 3015048414
             Email: tads@anri.barc.usda.gov
             Single pass sequencing. Bases called and trimmed with phred
             0.000925 using options -trim_alt " -trim fasta. Vector identified
             by cross_match using options -minmatch 12 -minscore 18
             Plate: 4 row: B column: 23
             Seq primer: CCCAGTCACGACGTGTAAAACG
             High quality sequence stop: 717.
FEATURES          Location/Qualifiers
     source          1..717
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /strain="Holstein"
                     /db_xref="taxon:9913"
                     /clone="9BOV4_E23"
                     /sex="Male"
                     /tissue_type="Pooled"
                     /dev_stage="Multiple"
                     /clone_lib="BARC 9BOV"
                     /lab_host="DH10B T1 phage resistant"
                     /note="Organ: Abomasum; Vector: pAgen-1; Site_1: EcoRV;
                     Site_2: NotI; Equimolar amounts of mRNA extracted from
                     fundic and pyloric abomasums of 18 and 21 week old steers.
                     Exposure to Ostertagia ostertagi was initiated at 15 weeks
                     of age. fundic and pyloric abomasum"
ORIGIN

Query Match          37.0%;          Score 526.4;          DB 7;          Length 717;
Best Local Similarity          83.7%;          Pred. No. 7.5e-124;
Matches 596;          Conservative 0;          Mismatches 116;          Indels 0;          Gaps 0;

Qy       250 CTCGGCCGGCGGTGCAGCATGACCCTGTGGGAACGGCGGCGGTACTGCCTTTTACCGCCC          309
             |||  |||                             ||  |||||||||||||||
Db         6 CTCAGCCGGGCTGCAGCATGACCTCGTGGGAATGGTGCCCTTCTACCCTCAGCCC           65

Qy       310 GGCATGCCGGCAGGGCTTCAGGGTTCCACTCGTCTCATCGTTATTCTAGTTTTTTGGGCTTCTAG          369
             |||  |||||  |||||||||||||||||||||||||||||||||||||||||
Db        66 GGCATCCCGGCCGCGGCCTTAGTCGTCCCCGCTACTCATCGTTCTTCTCGTTCTTGGTCTCTTGG          125

Qy       370 CAGCAAGCTTCCTGTCATCTTGCCGGGGGATCCGTTCGGCGAAAATGTGGCGTTCAGTGCAG          429
             |||||||||||||||||||||||||||||||||||||||||||||
Db       126 CCGCCAGCTTCCGTGCTCATCTTACCCGGGGGATTCGGTGCCCACTCGCGGCACTTCAGTGGTTGG          185

Qy       430 TGAGAGTTCTTCTCAGTCGTCGTTCAGTGAACACCAAACATCCTACAGAAAGCCCTTCAGCGGTGGAAGCGG          489
             ||||||||||||||||||||||||||||||||||||||||||||||
Db       186 TGAGAGTTCTTCTCAGCCGTCCTGTTCATCATAGGAGCAGGAGAGAAAATTGTCGGCCCGGGCACTTCAGCGAG          245

Qy       490 AATGGTTCGTTCGGGGTACAGTGGAACACCAACACACACATTCTCACACACACT          549
             ||||||||||||||||||||||||||||||||||||||||
Db       246 AATGGTTCAGTCGGCCACGGTTAGCACCAAAAACACTCCTAGGCGTTGGAAGCGCTTCAGCGGAGCCG          305

Qy       550 TTACAGCCCGTGTCGGTTCGTCGGTTCGGTCTCGCTCGCCCGCTTGGTCATCCCCCGTCGTGA          609
             ||||||||||||||||||||||||||||||||||||||||
Db       306 TCCGGACCCCACGTCGGTCTCGTCGTCATTGTGGGCTGGGCCGCCGTTAATATCACACTCACAGGGA          365

Qy       610 CCCCAGTGCATCAGCTGAACGAGACCAATTGACTACTACAACGAGACGGTTCACCTGGCGTCTGA          669
             ||||||||||||||||||||||||||||||||||||||||
Db       366 ACCCAGTGCACGAGCTGAAGCAGGTACGGCAACCGACCACTGAGACAGAGTGTGCCGGACCCAGTGC          425

Qy       670 AAGAGAATTACGGCCGGGGAGTACGGCGAACGACCACTGAGACAGAGTGTGCCGAGACCCAGTGC          729
             ||||||||||||||||||||||||||||||||||||||||
Db       426 GCCAAAACTATCCAGGGCGGTACCGCGAGGGCCTGGAGGAGGGCGGTTAATATCACACTCACAGGGA          485

Qy       730 TCTACCTGGGCGGGAGAAGTTCACACGGAGTAGCCTTGCGGCCTCCAACAGTAGCCACC          789
             ||||||||||||||||||||||||||||||||||||||||
Db       486 TCTAYCTGGGCGGGAGAAGTTCACTCCGAGCAGCCCCCTGGGCGTTTACCGCCAATACCGCC          545

Qy       790 TGGCGGGACACTACGGCCTCGGCCCACCGCTATGGGGTGGGGTTCTTGCCTTCTTGGGCGTTCCTCCA          849
             ||||||||||||||||||||||||||||||||||||||||
Db       546 TGGCGGGACACTACGGCTTCGGCCCACTCTATGGGTGGGCCCTTCGGCCTTCGCGTTCCTCCA          605

Qy       850 ACGTGTGCTCTCCAGCGCCCCCCGCTCTACGGGAGCCCTGGCGACTGCGCACCACCGGAG          909
             ||||||||||||||||||||||||||||||||||||||||
Db       606 ACATGCTGCTCTCCATCGCGGTTCCGGCACTAGGGAGCCCTGACTCTCTCATCACCGGCG          665

Qy       910 CCTTTGCGGGCTCTTCGGGGTCTTCGCCTTGGGCCTTCCATCTCTAGGGGTGCGGCG          961
             |||||||||||||||||||||||||||||||||||||||
Db       666 CCTTCGCGGGCTCTTCTCGGTGTTCGCCGTGTTCGCCCTCTAYCTCACGGGCCTTCCTCT          717

RESULT 11
CK979918
LOCUS        CK979918          703 bp          mRNA          linear          EST 16-MAR-2004
DEFINITION   4111631 BARC 9BOV Bos taurus cDNA clone 9BOV40_D22 5', mRNA
```

(left column — alignment for RESULT 10 / CK971974)

```
Qy        57 CCAGGGAAAGTCTCCTGGGGCGATCGTGTAAATAAAACCTTTTTTTCTTTTGTTTTTT 2          1421
             ||||
Db        57 CCAGGGAAAGTCTCCTGTGTAAAATAAAACCTTTTTTTCTTTTATGTATTT 2          58

Qy       117 GGAGAGCTTCAGGAAGGGCACTGAGCGCTGCCGGGGACCTCGGACAATCCCGACGAGGCA          117
             ||||
Db       117 ...          58

Qy       177 ACTCCTTCCCCGCCGCCTTGGGACATGGCGCAGGGGGAAGCAGTAGCAGTGCGGCGGCCA          118
             ||||||||||||||||||||||||||||||||||||
Db       177 ACTCCTTCCCCGCCGCCTTAGTCGTCCCCGCTACTCATCGTTCTTCTCGTTCTTGGT          118

Qy       237 AGGCGCGCTCTCCCAGACTTAAAAATGTATCACCACTAACCTGTGAGGGGACCCAATCTGG          237
             |||||||
Db       237 AGGCGCTCTCCCAGACTTAAAAATGTATCACCACTAACCTGTGAGGGGACCCAATCTGG          178

Qy      1186 AGGCCGCGCTCTCCCAGACTTAAAAATGTATCACCACTAACCTGTGAGGGGACCCAATCTGG          1245
             |||||||||||||||||||||||||||||||||||||
Db      1186 AGGCCGCGCTCCCAGACTTAAAAATGTATCACCACTA          1245

Qy      1126 AGGACTGCAGCCACGAGAGAGGGGGCTCACCTCTTATCTCTCGGCCGCAGCCCACTGCACAAGC          1185
             ||||||||||||||||||||||||||||||||||||||
Db      1126 AGGACTGCAGCCACGAGAGAGGGGGCTCACCTCTTATCTCTCGGCCGCAGCCCACTGCACAAGC          1185

Qy      1306 GGAGAGCTTCAGGAAGGGCACTGAGCGCTGCCGGGGACCTCGGACAATCCCGACGAGGCA          1365
             ||||||||||||||||||||||||||||||||||||
Db      1306 GGAGAGCTTCAGGAAGGGCACTGAGCGCTGCCGGGGACCTCGGACAATCCCGACGAGGCA          1365

Qy      1246 ACTCCTTCCCCGCCGCCTTGGGACATGGCGCAGGGGGAAGCAGTAGCAGTGCGGCGGCCA          1305
             ||||||||||||||||||||||||||||||||||||||||
Db      1246 ACTCCTTCCCCGCCGCCTTGGGACATGGCGCAGGGGGAAGCAGTAGCAGTGCGGCGGCCA          1305

Qy       297 AGGACTGCAGCCACGAGAGAGGGGGCTCACCTCTTATCTCTCGGCCGCAGCCCACTGCACAAGC          238
             ||||||||||||||||||||||||||||||||||||||
Db       297 AGGACTGCAGCCACGAGAGAGGGGGCTCACCTCTTATCTCTCGGCCGCAGCCCACTGCACAAGC          238

Qy       357 TGGTGAGTCTCCAGTATGTTCGGCCCCAGGGCTCTTCGCACCCTTCTGGACCAAAGCGGCCA          298
             |||||||||||||||||||||||||||||||||||||
Db       357 TGGTGAGTCTCCAGTATGTTCGGCCCCAGGGCTCTTCGCACCCTTCTGGACCAAAGCGGCCA          298

Qy      1066 TGGTGAGTCTCCAGTATGTTCGGCCCCAGGGCTCTTCGCACCCTTCTGGACCAAAGCGGCCA          1125
             ||||||||||||||||||||||||||||||||||||||
Db      1066 TGGTGAGTCTCCAGTATGTTCGGCCCCAGGGCTCTTCGCACCCTTCTGGACCAAAGCGGCCA          1125

Qy       417 GCGCGGCCTTCGGGTCACGGCTGGCAACCGGGCGTCCTGTGGCCTCCTCGCGGGTCTCACCACTCAGTACG          358
             |||||||||||||||||||||||||||||||||||||
Db       417 GCGCGGCCTTCGGGTCACGGCTGGCAACCGGGCGTCCTGTGGCCTCCTCGCGGGTCTCACCACTCAGTACG          358

Qy      1006 GCGCGGCCTTCGGGTCACGGCTGGCAACCGGGCGTCCTGTGGCCTCCTCGCGGGTCTCACCACTCAGTACG          1065
             ||||||||||||||||||||||||||||||||||||||
Db      1006 GCGCGGCCTTCGGGTCACGGCTGGCAACCGGGCGTCCTGTGGCCTCCTCGCGGGTCTCACCACTCAGTACG          1065

Qy       477 TCTCTAGCGGTGCCCGCTCTCGCCCGCTCCGGCGCTCCCGGCGGTTCCGCGCTCTCCGGGCGTTCGCCTTGGCCTTCCA          418
             |||||||||||||||||||||||||||||||||||||
Db       477 TCTCTAGCGGTGCCCGCTCTCGCCCGCTCCGGCGCTCCCGGCGGTTCCGCGCTCTCCGGGCGTTCGCCTTGGCCTTCCA          418

Qy       946 TCTCTAGCGGTGCCCGCTCTCGCCCGCTCCGGCGCTCCCGGCGGTTCCGCGCTTCGCCTTGGCCTTCCA          1005
             ||||||||||||||||||||||||||||||||||||||
Db       946 TCTCTAGCGGTGCCCGCTCTCGCCCGCTCCGGCGCTCCCGGCGGTTCCGCGCTTCGCCTTGGCCTTCCA          1005

Qy       537 GCCTGCGCACTGCGTGAACACCGGAGCCTTCGGCGCCTTCGGCCGCCTTCGGCCGCTTCGGCCTCCA          478
             |||||||||||||||||||||||||||||||||||||
Db       537 GCCTGCGCACTGCGTGAACACCGGAGCCTTCGGCGCCTTCGGCCGCCTTCGGCCGCTTCGGCCTCCA          478

Qy       886 GCCTGGCACTGCTGACCACCGGAGCCTTGCGCTTCGGGTGTTCTTGGGGTTCGGGTTCGGGTTCTTGGCCTTCCA          945
             ||||||||||||||||||||||||||||||||||||||
Db       886 GCCTGGCACTGCTGACCACCGGAGCCTTGCGCTTCGGGTGTTCTTGGGGTTCGGGTTCGGGTTCTTGGCCTTCCA          945
```

```
Sequence.
ACCESSION   CK979918
VERSION     CK979918.1  GI:45497898
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 703)
AUTHORS     Sonstegard, T.S., Van Tassell, C.P., Matukumalli, L.K., Harhay,
            G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
TITLE       Production of EST from cDNA libraries derived from immunologically
            activated bovine gut
JOURNAL     Unpublished (2004)
COMMENT     Contact: Tad S. Sonstegard
            Bovine Functional Genomics Laboratory
            Animal and Natural Resources Institute
            Bldg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
            Tel: 3015048416
            Fax: 3015048414
            Email: tads@anri.barc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            0.000925 using options -trim_alt "-trim_alt" .fasta. Vector identified
            by cross_match using options -minmatch 12 -minscore 18
            Plate: 40 row: D column: 22
            Seq primer: CCCAGTCACGACGTTGTAAAACG
            High quality sequence stop: 703.
FEATURES             Location/Qualifiers
     source          1..703
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /strain="Holstein"
                     /db_xref="taxon:9913"
                     /clone="9BOV40_D22"
                     /sex="Male"
                     /tissue_type="Pooled"
                     /dev_stage="Multiple"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="BARC 9BOV"
                     /note="Organ: Abomasum; Vector: pAgen-1; Site 1: EcoRV;
                     Site 2: NotI; Equimolar amounts of mRNA extracted from
                     fundic and pyloric abomasums of 18 and 21 week old steers.
                     Exposure to Ostertagia ostertagi was initiated at 15 weeks
                     of age. fundic and pyloric abomasum"
ORIGIN

Query Match      36.7%;  Score 522.2; DB 7; Length 703;
Best Local Similarity   83.9%;  Pred. No. 9e-123;
Matches   590;  Conservative    0; Mismatches  113; Indels    0; Gaps    0;

Qy    288 CTGCCTTTTTACCCCCAGCCCCGGCATGCCGCAGGCTTCAGCGTTCCACTGCTCATCGTT 347
          |||||| ||| ||||||||||||||| ||||||| |||||||||||||||||||
Db      1 CTGCCCTTCACCCTCAGCCCCGCATGCCGCGGCCTTAGTGTCCCGCTACTCATCGTC 60

Qy    348 ATTCTAGTGTTTTTTGGGCTCTAGCAGCAGCTTCCTGCTCATCTTGCCGGGGATCCGTGGC 407
          |||  |||||  |||||  | |||  |||||||||||||||||| || || || |||||||
Db     61 ATTCTCGTGTTCTTTGGGTCTTCGGCGCCAGCTTCCTGCTCATCTTACCCGGGATTCGTGGC 120

Qy    408 CACTCGCGCCTCGGTTTTGGTTGGTGAGAGTTCTTCTCAGTTCATTCATAGGCCAGAAATT 467
          ||||| || |||||  || |||  ||||| ||||||| ||||| ||| |||||||||||
Db    121 CACTCGCGAATGGTTTCTCGGTTGGTGGAGAGTTCTTCTCAGCCTGTTCATAGGAGCAGAAATT 180

Qy    468 GTGGCGTGTGCCACTTCAGTGACGAATGGGTTCGTGGGTACAGTGACACCAACACATCCTAC 527
          ||||| ||| || | ||||||  ||  |  || || || | || | | || ||||||||||
Db    181 GTGGCCGCGGCACTTCAGGCGACGAATGGTTCAGTCGGGCAGCGGTTAGCACCAAAACATCCTAC 240

Qy    528 AAAGCCTTCAGCGGCAGGGGCCGTTACAGCCCCGTTGGTGGTTCGGGCGGCCTGGAGGGC 587
          ||  |||||| ||||   ||| ||| ||||||||| ||| || ||||| | |||||||||
Db    241 AAGGCCTTCAGCGCTGGAAGCGCGTCCGGACGTCCGTCCATGTGGGCCTGAGGGC 300

Qy    588 ATTAATAATTACACTCACAGGGAACCCCAGTGCATCAGGTGAACGAGAGACCATTGACTACAC 647
          ||  | || || |||||||| |  | |  ||||||||||  |  ||  | ||| ||||||||
Db    301 GTTAATTATCACACTCACAGGGAACCACACGTGCAGCGGAGCTGAACGAGAACGCATTGACTACAAT 360

Qy    648 GAGCAGTTCACCTGCGTCTGAAAAGAAGATTACCGCGCGGGAGTGACGGTACGGGAAACGCACTGAG 707
          |||||||||| |||| | || |  | |||| |||||| ||| ||  || | || ||||| |||||
Db    361 GAGCAGTTCATCTGCGCGTTTGGCCAAAACTATGCCAGGGGCGTAGCGGCCCCTGGAG 420

Qy    708 AAGGGGCTGCCGGACCCAGTGCTCTACCTGGGCGGGAGGAGTTCACCAGTAGCCCTTGC 767
          ||| || |||||||||||||| |||| |||| || |||||||||||| || || | |||
Db    421 AGGGGGCTGCCGGACCCGGTTCTCTATCTGGCGGAGAGTTCACTCCGAGCAGCCCTGT 480

Qy    768 GGCCTGTACCACCAGTTACCACTTGGCGGGGAGCACTACCGCCACGGGATCTATGGGTGGCGG 827
          |||| ||| ||| | |  | || |  |  || ||||| || ||| || | |||| | ||
Db    481 GGGGTTTACCGCCAATACCGGCTGGGCGACGAGCACTACGGCCAGTCTATGGGTGGCC 540

Qy    828 TTCTGCTTCTGGGTCTCCTCTCCAGCGTGGTGCTCTCCAGCGCGGCGCCCCGTCTACGGAGGC 887
          ||||| |||  | || ||||||||  || | || ||||| | |||||| || | ||||| |||
Db    541 TTCTGCTTCTGGGTCTCCTCTCCAACATGCTGCTCTCCATGCGGTTCACAGGCCGCGCTACGGGAGGC 600

Qy    888 CTGGCACTGCTGACCACCGGAGCCTTCGGGTCTTCGGGGTCTTCGGCCTTGGGCCTTCATC 947
          ||||||| || ||| ||||| |||||||||  || || | || || |||||| |||||||
Db    601 CTGACTCCTCCTCATCACTGGCGCGTCTCTTCGGCGGGTGTTGCGCCTTTCGGCCTCTATC 660

Qy    948 TCTAGGCGTGCCGCTCTCGCCGCGTCCCCGCTCCGCCCTTAGGGTCTTCCTCCCGCGC 990
          |||| |  |||  |  |  |||  |  || || |||||||| || | |||||| |||
Db    661 TCCAGCCGTGCCCTCTTCGCCAGACCTGCGTGCGGTCGGGCTTCGCCGTCTATC 703
```

ORIGIN

I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

```
Query Match      36.0%;   Score 512; DB 1; Length 515;
Best Local Similarity 100.0%;   Pred. No. 3.6e-120;
Matches   512;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   910 CCTTCGCGCTCTTCGGGGTCTTCGGCGGTGCCCATCTCTAGCGGTGCGCGTCTGCCCCGC  969
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   515 CCTTCGCGCTCTTCGGGGTCTTCGGCGGTGCCCATCTCTAGCGGTGCGCGTCTGCCCCGC  456

Qy   970 TCCGGCTAGGCTCCTCCGCGCTCACCCACTCAGTACGGGCGGCGCCGCCTTCTGGGTCAGCGCTGG  1029
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   455 TCCGGCTAGGCTCCTCCGCGCTCACCCACTCAGTACGGGCGGCGCCGCCTTCTGGGTCAGCGCTGG  396

Qy  1030 CAACCGGCCGTCCTGGCCTTCTCTCCGAGGGGGGCGCCGTGGTGAGTCTCAGTATGTTCGGC  1089
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   395 CAACCGGCCGTCCTGTGCCTTCTCTCGGAGGGGGGCCGTGGTGAGTCTCCAGTATGTTCGGC  336

Qy  1090 CCAGGCGCTCTTCGCACCCTTCTCGGACCAAAGCGCCAAGGACTGCAGCGAGGAGGAGGGG  1149
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   335 CCAGCGCTCTTCGCACCCTTCTCGGACCAAAGCGCCAAGGACTGCAGCGAGGAGGAGGGG  276

Qy  1150 GCTCACCCTTTATCCTCGGCGGACCCACTGCACAAGCAGCGCCGCTCTCCCAGACTTAAAAT  1209
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   275 GCTCACCCTTTATCCTCGGCGGACCCACTGCACAAGCAGCGCCGCTCTCCCAGACTTAAAAT  216

Qy  1210 GTATCACCACTAACCTGTGAGGGGGGACCCAAATCTCGGACTCCTTCCCCGCCTTGGGACATC  1269
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   215 GTATCACCACTAACCTGTGAGGGGGGACCCAAATCTCGGACTCCTTCCCCGCCTTGGGACATC  156

Qy  1270 GCAGGCCCGGGAAGCAGTGCCCGCGCCAGGGGCTCCAGGAAGGGCCACTGA  1329
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   155 GCAGGCCCGGGAAGCAGTGCCCGCGCCAGGGGCTCCAGGAAGGGCCACTGA  96

Qy  1330 GCGCTGCTGGCGTCGGGACATCCGCGAGGCCGGGAAGTCTCCTGGGCGGAT  1389
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    95 GCGCTGCTGGCGTCGGGACATCCGCGAGGCCGGGAAAGTCTCCTGGGCGGAT  36

Qy  1390 CTGTAAATAAACCTTTTTTTCTTTTGTTTTTT  1421
         ||||||||||||||||||||||||||
Db    35 CTGTAAATAAACCTTTTTTTCTTTTGTTTTTT  4
```

RESULT 13
AI304327/c

LOCUS   AI304327   510 bp   mRNA   linear   EST 01-FEB-1999
DEFINITION   qo57h06.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1912667 3',
             similar to contains TAR1.t3 TAR1 repetitive element ;, mRNA
             sequence.
ACCESSION   AI304327
VERSION   AI304327.1   GI:3988016
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominidae; Homo.
REFERENCE   1 (bases 1 to 510)
AUTHORS   Robert Strausberg, Ph.D.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
             Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
             Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be

---

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1487   Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 476.

FEATURES   Location/Qualifiers
     source   1..510
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:1912667"
                  /tissue_type="adenocarcinoma"
                  /clone_lib="NCI_CGAP_Co8"
                  /lab_host="DH10B"
                  /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                  modified polylinker; 1st strand cDNA was prepared from
                  colon adenocarcinoma, and was then primed with a Not I -
                  oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                  RI adaptors (Pharmacia), digested with Not I and cloned
                  into the Not I and Eco RI sites of the modified pT7T3
                  vector. Library is normalized. Library was constructed by
                  Bento Soares and M. Fatima Bonaldo. "

ORIGIN

```
Query Match      35.9%;   Score 510; DB 1; Length 510;
Best Local Similarity 100.0%;   Pred. No. 1.2e-119;
Matches   510;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   912 TTCGCGCTCTTCGGGGTCTTCGGCGGTGCCCATCTCTAGCGGTGCGCGTCTGCCCCGCTC  971
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   510 TTCGCGCTCTTCGGGGTCTTCGGCGGTGCCCATCTCTAGCGGTGCGCGTCTGCCCCGCTC  451

Qy   972 CGGCTAGGCTCCTCCGCGCTCACCCACTCAGTACGGGCGGCGCCGCCTTCTGGGTCAGCGCTGGCA  1031
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   450 CGGCTAGGCTCCTCCGCGCTCACCCACTCAGTACGGGCGGCGCCGCCTTCTGGGTCAGCGCTGGCA  391

Qy  1032 ACCGGCCGTCCTGTGCCTTCTCTCCGGAGGGGGCCGTGGTGAGTCTCCAGTATGTTCGGCCCC  1091
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   390 ACCGGCCGTCCTGTGCCTTCTCTCCGGAGGGGGCCGTGGTGAGTCTCCAGTATGTTCGGCCCC  331

Qy  1092 AGCGCTCTTCGCACCCTTCTCGGACCAAAGCGCCAAGGACTGCAGCGAGGAGGAGGGGGGC  1151
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   330 AGCGCTCTTCGCACCCTTCTCGGACCAAAGCGCCAAGGACTGCAGCGAGGAGGAGGGGGGC  271

Qy  1152 TCACCCTTTATCCTCGGCGGACCCACTGCACAAGCAGCGCCGCTCTCCCAGACTTAAAATGT  1211
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   270 TCACCCTTTATCCTCGGCGGACCCACTGCACAAGCAGCGCCGCTCTCCCAGACTTAAAATGT  211

Qy  1212 ATCACCACTAACCTGTGAGGGGGGACCCAAATCTCGGACTCCTTCCCCGCCTTGGGACATCGC  1271
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   210 ATCACCACTAACCTGTGAGGGGGGACCCAAATCTCGGACTCCTTCCCCGCCTTGGGACATCGC  151

Qy  1272 AGGCCCGGGAAGCAGTGCCCGCGCCAGGGGCTCCAGGAAGGGCCACTGAGCGCTGCTGGCGTCAGC  1331
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   150 AGGCCCGGGAAGCAGTGCCCGCGCCAGGGGCTCCAGGAAGGGCCACTGAGCGCTGCTGGCGTCAGC  91

Qy  1332 GCTGCTGGCGGCCAGGGCTTCGGACATCCGCGACACCAGGGGAGAAAGTCTCCTGGGCGGATCT  1391
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    90 GCTGCTGGCGGCCAGGGCTTCGGACATCCGCGACACCAGGGGAGAAAGTCTCCTGGGCGGATCT  31

Qy  1392 GTAAATAAACCTTTTTTTCTTTTGTTTTTT  1421
         ||||||||||||||||||||||||||||||
Db    30 GTAAATAAACCTTTTTTTCTTTTGTTTTTT  1
```

RESULT 14
BP396416

LOCUS   BP396416   500 bp   mRNA   linear   EST 03-FEB-2005
DEFINITION   BP396416 Homo sapiens pancreatic islet Homo sapiens cDNA clone
             hbi10023 3', mRNA sequence.
ACCESSION   BP396416
VERSION   BP396416.1   GI:58569059
KEYWORDS   EST.

```
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominidae; Homo.
REFERENCE    1 (bases 1 to 500)
  AUTHORS    Takeda,J., Jin,L. and Horikawa,Y.
  TITLE      Expression profile of mRNAs from human pancreatic islet tumors
  JOURNAL    Unpublished (2005)
COMMENT      Contact: Yukio Horikawa
             Laboratory of Molecular Genetics
             Institute for Molecular and Cellular Regulation, Gunma University
             Showa-machi 3-39-15, Maebashi, Gunma, 371-8512, Japan
             Tel: 81-27-220-8832
             Fax: 81-27-220-8889
             Email: yhorikaw@showa.gunma-u.ac.jp.
FEATURES     Location/Qualifiers
  source     1..500
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="hbt10023"
             /tissue_type="pancreatic islet"
             /clone_lib="Homo sapiens pancreatic islet"
ORIGIN

Query Match            34.9%;  Score 496;  DB 3;  Length 500;
Best Local Similarity 100.0%;  Pred. No. 4.5e-116;
Matches  496;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    926 GGTGCTTCGGCTTGGCCTTCCATCTCTAGCGTGCGGCGGCTCTTGCCCCGTTCCGCCTAGGCTCTTC 985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGTGCTTCGGCTTGGCCTTCCATCTCTAGCGTGCGGCGGCTCTTGCCCCGTTCCGCCTAGGCTCTTC 60

Qy    986 CGCGGCTCACCACTCAGTACGGGCGCCGCCCGGCCTTCTGGGTCAGGCTGGCAACCGGCGCCTGTG 1045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCGGCTCACCACTCAGTACGGGCGCCGCCCGGCCTTCTGGGTCAGGCTGGCAACCGGCGCCTGTG 120

Qy   1046 CCCTCTTCCTCGGAGGGGCCGTGGTCAGTCTCCAGTATGTTTCGGCCCCAGGCGTCTTGCGAC 1105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCCTCTTCCTCGGAGGGGCCGTGGTCAGTCTCCAGTATGTTTCGGCCCCAGGCGTCTTGCGAC 180

Qy   1106 CCTTCTCTGACCAAAGCGCCAAGGACTGCAGCCAGGAGAGAGGGGCTCACCTCTTATCCT 1165
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTTCTCTGACCAAAGCGCCAAGGACTGCAGCCAGGAGAGAGGGGCTCACCTCTTATCCT 240

Qy   1166 CGGGCGACCCACTGCACAAGCCAGCCAGGAGCCGCTCTCCCAGACTTAAAAATGTATCACC... 1225
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGGCGACCCACTGCACAAGCCAGCCAGGAGCCGCTCTCCCAGACTTAAAAATGTATCACC... 300

Qy   1226 GTGAGGGGGACCCAATCTGGACTCTCTTCCCGCGGCCCTTGGGACATCGCCAGGCGCGGGAGCAG 1285
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GTGAGGGGGACCCAATCTGGACTCTCTTCCCGCGGCCCTTGGGACATCGCCAGGCGCGGGAGCAG 360

Qy   1286 TGCCCGGCCAGGCCTTGGGCCTCAGGGAAGGGGCACTGCGTGCGTGGCGCGGAG 1345
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGCCCGGCCAGGCCTTGGGCCTCAGGGAAGGGGCACTGCGTGCGTGGCGCGGAG 420

Qy   1346 GCCTTCGGACATCCCGCGGGGCACCGGGGAAAGTCTCCTGGGGCGGACTGTGTAAATAAAACCTTT 1405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCTTCGGACATCCCGCGGGGCACCGGGGAAAGTCTCCTGGGGCGGACTGTGTAAATAAAACCTTT 480

Qy   1406 TTTTCTTTTTTGTTTTTTT 1421
          |||||||||||||||||||
Db    481 TTTTTCTTTTTTGTTTTTTT 496
```

```
VERSION     AI791844.1  GI:5339486
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 505)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
COMMENT     Other ESTs: nk08a11.s1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
            Ph.D.
            cDNA Library Arraying: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            This read is a RESEQUENCE of a previously sequenced human clone.
            Original clone citation: see original entry for original citation
            information.
            This 5' resequenced clone has no previous 5' data to verify this
            new read against.
            Putative full length read
            The vector to vector length read
            Insert length: 741    Std Error: 0.00
            Seq primer: -40RP from Gibco
            High quality sequence stop: 418.
FEATURES    Location/Qualifiers
  source    1..505
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:1012892"
            /tissue_type="tumor"
            /lab_host="SOLR (kanamycin resistant)"
            /clone_lib="NCI_CGAP_Co2"
            /note="Organ: colon; Vector: Bluescript SK-; Site 1:
            EcoRI; Site 2: XhoI; Cloned unidirectionally.
            Oligo dT. Bulk colon villous adenoma. 5' adaptor sequence:
            5' GAATTCGGCACGAG 3'; 3' adaptor sequence: 5'
            CTCGAGTTTTTTTTTTTTTTTTTTT 3' Average insert size: 1.1 kb."
ORIGIN

Query Match            34.4%;  Score 489.2;  DB 1;  Length 505;
Best Local Similarity 99.2%;  Pred. No. 2.5e-114;
Matches  502;  Conservative  0;  Mismatches  3;  Indels  1;  Gaps  1;

Qy    744 AAGTTCACACCGAGTAGCCCTTGCGGCCGCCTGTACCACAGTACCACCTGGCGCGGGACAACTAC 803
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAGTTCACACCGAGTAGCCCTTGCGGCCGCCTGTACCACAGTACCACCTGGCGCGGGACAACTAC 60

Qy    804 GCCTTCGGGCCACGCTATGGGGTGGGCGTTCTTCGGCTCTCTTCTCAACGTGCGTGCTCTCC 863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCCTTCGGGCCACGCTATGGGGTGGGCGTTCTTCGGCTCTCTTCTCAACGTGCGTGCTCTCC 120

Qy    864 ACGGCCGGGCCCCCGGCTCTACGGAGGCCTGGCCACTGCTGACCACCGAGGACCTTCGGCGTTCTTC 923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACGGCCGGGCCCCCGGCTCTACGGAGGCCTGGCCACTGCTGACCACCGAGGACCTTCGGCGTTCTTC 180

Qy    924 GGGGGTCTTCGGCCTTGGCCTCCATCATCTCTAGCGTGCCGCTCTGCCCCGTTCCGCCTAGGCTCC 983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGGGGTCTTCGGCCTTGGCCTCCATCATCTCTAGCGTGCCGCTCTGCCCCGTTCCGCCTAGGCTCC 240

Qy    984 TCCGCGCGCTCACCACTCAGTACGGGCGCCGCCCGGCCTTCTGGGTCAGGCTGGCAACCGGCGTCCTG 1043
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ...
```

```
Db  241 TCCGCGCTCACCACTCAGTACGGACGCGCCCGCCTTCTGGGGTCACGCTGGCAACCGGCGTCCTG 300

Qy 1044 TGCCTCTTCCTCGGAGGGGCCGTCGGTGAGTCTCCAGTATGTTCGGCCCAGGCGCTCTTCGC 1103
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 TGCCTCTTCCTCGGAGGGGGCCGGTGGTGAGTCTCCAGTATGTTCGGCCCAGGCGCTCTTCGC 360

Qy 1104 ACCCTTCTTGGACCAAAGCGCCCAAGGACTTGCAGCCAGGAGGAGGGGGCTTCACCTTCTTATC 1163
            ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ACCCTTCTTGGACCAAAGCGCCCAAGGACTCGGCACCAGGAGGAGGGGGCTCACCTTCTTATC 420

Qy 1164 CTCGGGCGACCCACTGCACAAGCAGGAGCGCCGGCTTCTCCCAGACTTAAAAGTTATCACCACTAAC 1223
            |||||||||||||||||||||||||||||||||||||||  |||||||||||||||||||||
Db  421 CTCGGGCGACCCACTGCACAAGCAGGAGCGCCGGCTCTCCCAGACTT-AAATGTTATCACCACTAAC 479

Qy 1224 CTGTGAGGGGGACCCAATCTGGGACTC 1249
            |||||||||||||||||||||||||||
Db  480 CTGTGAGGGGGACCCAATCTGGGACTC 505
```

Search completed: January 22, 2006, 02:34:11
Job time : 6255 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2006 Compugen Ltd.

GenCore version 5.1.6

Run on: January 21, 2006, 22:22:21 ; Search time 286 Seconds
(without alignments)
8831.871 Million cell updates/sec

Title: US-09-937-059-57
Perfect score: 1421
Sequence: 1 gaaccaggaaagtaacggct..........cttttttctttgtttctt 1421

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	354.8	25.0	1029	3	US-09-578-063-54	Sequence 54, Appl
2	354.8	25.0	2133	3	US-09-578-063-53	Sequence 53, Appl
3	61	4.3	7218	2	US-08-232-463-14	Sequence 14, Appl
4	50.2	3.5	1098	3	US-09-252-991A-1304	Sequence 1304, Ap
5	50.2	3.5	1257	3	US-09-252-991A-1255	Sequence 1255, Ap
6	50.2	3.5	1746	3	US-09-252-991A-1163	Sequence 1163, Ap
7	48.2	3.4	396	3	US-09-252-991A-1350	Sequence 1350, Ap
8	48	3.3	1436	3	US-09-614-912-35	Sequence 35, Appl
9	47.6	3.3	420	3	US-09-252-991A-11029	Sequence 11029, A
10	47.6	3.3	1683	3	US-09-252-991A-11226	Sequence 11226, A
11	47.6	3.3	1974	3	US-09-252-991A-11111	Sequence 11111, A
12	47.6	3.3	2283	3	US-09-252-991A-10956	Sequence 10956, A
13	46.4	3.3	1431	3	US-09-614-912-49	Sequence 49, Appl
14	46	3.2	601	3	US-09-949-016-88511	Sequence 88511, A
15	46	3.2	8991	3	US-09-949-016-14248	Sequence 14248, A
16	45.6	3.2	1037	3	US-09-902-540-6939	Sequence 6939, Ap
17	45.6	3.2	2283	3	US-09-252-991A-10956	Sequence 599, App
18	45.2	3.2	3521	3	US-09-902-540-599	Sequence 599, App
19	45	3.2	15840	3	US-09-902-540-1134	Sequence 1134, Ap
20	45	3.2	1926	3	US-09-249-585A-4	Sequence 2, Appli
21	44.8	3.2	1931	2	US-09-130-114-2	Sequence 2, Appli
22	44.8	3.2	2229	3	US-09-902-540-4193	Sequence 4193, Ap
23	44.4	3.1	21330	3	US-09-902-540-1209	Sequence 1209, Ap
24	44.4	3.1	2901	3	US-09-949-016-4707	Sequence 4707, Ap
25	44.4	3.1	3045	3	US-09-949-016-701	Sequence 701, App
26	44.4	3.1	9506	3	US-09-949-016-16449	Sequence 16449, A
27	44.4	3.1	11858	3	US-09-949-016-12443	Sequence 12443, A
28	44.4	3.1	152331	3	US-09-128-155-16	Sequence 16, Appl
29	43.8	3.1	462	3	US-09-252-991A-1610	Sequence 1610, Ap
30	43.8	3.1	1362	3	US-09-252-991A-1556	Sequence 1556, Ap
31	43.8	3.1	1551	3	US-09-252-991A-12931	Sequence 12931, A
32	43.8	3.1	1587	3	US-09-252-991A-1561	Sequence 1561, Ap
33	43.8	3.1	1659	3	US-09-252-991A-13081	Sequence 13081, A
34	43.8	3.1	1917	3	US-09-252-991A-1465	Sequence 1465, Ap
35	43.8	3.1	2073	3	US-09-252-991A-12858	Sequence 12858, A
36	43.4	3.1	597	3	US-09-252-991A-5066	Sequence 5066, Ap
37	43.4	3.1	1308	3	US-09-252-991A-5127	Sequence 5127, Ap
38	43.4	3.1	1782	3	US-09-252-991A-5094	Sequence 5094, Ap
39	43.2	3.0	4131	3	US-09-252-991A-13773	Sequence 13773, A
40	43.2	3.0	8211	3	US-09-252-991A-13656	Sequence 13656, A
41	43.2	3.0	58857	3	US-09-477-962-1	Sequence 1, Appli
42	43	3.0	699	3	US-09-821-255-1	Sequence 1, Appli
43	43	3.0	762	3	US-10-202-676-3	Sequence 3, Appli
44	43	3.0	835	3	US-10-202-676-1	Sequence 1, Appli
45	43	3.0	835	9	5223425-7	Patent No. 5223425

ALIGNMENTS

RESULT 1
US-09-578-063-54
Sequence 54, Application US/09578063

; Patent No. 6764677
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A
; APPLICANT: Barnes, Thomas M
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 210147.0023/6U1
; CURRENT APPLICATION NUMBER: US/09/578,063
; CURRENT FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 54
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-578-063-54

Query Match 25.0%; Score 354.8; DB 3; Length 1029;
Best Local Similarity 64.5%; Pred. No. 7.7e-78;
Matches 546; Conservative 0; Mismatches 297; Indels 3; Gaps 1;

Qy 287 ACTGCGCTTTTTACCCCAGCCCGGCCATGCCGCAGGCTTCACGGTTCCACTGCTCATCGT 346
        ||||| ||||||||||||| || ||||||||| | |||||||||||||| |||| ||||||
Db 21 ATTCCCCTTCTATAGCTGGCCCCCAAGCCAACCTTCCCGATGGACACCACTTTGGCCAGCAT 80

Qy 347 TATTCTAGTGTTTTTTGGCCTCTAGCAGAGAAGCTTCCTGCTCATCTTGCGCGGGATCCGTGG 406
        ||||||||||||| |||||| ||| | |||||||| |||||||||||||||| |||||||||
Db 81 CATCATGATCTTTCTGACTGCACTGGGCCACGTTCATCGTTCATCCCTGCCTGGGCATTCGGG 140

Qy 407 CCACTCCGCGCTCGTTTTTCGGTTGGTTGGAGAGTTCTTCTCAGTCTGTTCATAGGCGGCAGAAAT 466
        |||||||||||||| ||||| |||||| |||| |||||| ||||||||| |||| |||||| ||||
Db 141 AAAGACGGAGGCTGTTGCTTCAGTTGCTTCCGGGTGGTGACCAGCTTAATTCATCGGGCGCCAAT 200

Qy 467 TCTGGGCTGTGCACTTCAGTGCAGAATGGTTCTTCGTGGGTACAGTGAACAGCACCAACACATCCTA 526
        ||||||| |||||||||||||| |||||||| || |||| |||||||| ||| ||| |||||||||
Db 201 CCTGGGCTGTGAATTTCAGTTCTGAGTGGTTCTGTGGGCCAGGTCAGCAGCACCAACACATCATA 260

Qy 527 CAAAGCCTTCAGCGCAGCGCAGCGCGCCGTTACAGCCCGTGTCCGTCTCGTGGGCCTTGGAGGG 586
        ||||||||||||||||| || ||||||||||||| |||||||| |||||||||||||||||||||
Db 261 CAAGGCCTTCAGTTCTGAGTGGAATCAGCGCCCGTGATTATTCGGCGTCCAGGTCGGGCGTCGG 320

```
Qy    587  CATTAATAATTACACTCACAGGGACGACCCCAGTTGCATCAGCTGAACGAGACCAATTGACTACAA  646
           ||||||        ||||||||||          ||||| |||||||    |||||||||   |||||
Db    321  AGTCAACACTCACAGGGACCCCTGCACGAGCTGAATGAAGACCAATCAATTACAA  380

Qy    647  CGAGCAGTTCACCTGGCGTTCTGAAAGAGAATTACCGCGGGGAGTACGCGAACGCACTGGA  706
           ||||| ||||||||||||  |||||       ||    |||| |||   ||| |||||||
Db    381  CGAGGAGTTCACCTGGCGCCTGGAGAACTATGCTGAGGAGTGTGCAAAGGCTCTGGA  440

Qy    707  GAAGGGGCTGCCGGACCCAGTGCTTCTACCTGGGGGAGGAGAGTTCACACGAGCCCTTG  766
           |||||||||||||||||||||||||||||||    ||||||||||||||||||||||||
Db    441  GAAGGGGCTGCCGGACCCAGTGCTTGTTACCTTAGCTGAGAAGTTCACTCAGAAGCCCATG  500

Qy    767  CGGCCTGTACCACCAGTACCACCTGGCGGCGGCGGCACTACGCCCTATAGGGTGGC  826
           |||||||||||||||||||||    |||||||    |||||||||| ||||||||
Db    501  TGGCCTATACCGGCCAGTACCGGCTGGGCGGGCACTACCATCCTATGGGGTGGC  560

Qy    827  GTTCTGCTTCTGGCTGCTTCTCCAAGCGTGTTGTGCTCCAGCGCCCGGCGCCCTTACGGAGG  886
           ||||||||| |||||||||||||||||||  ||||  ||||||||||||    ||||||||
Db    561  ATTCCTCTTGGGCTGCGTCGGCGGCAAATGATGCTCTCCATGCCCTGTGTCTGGTAATGGTGG  620

Qy    887  CCTGGCACTGCTGACCACCGGGAGCCCTTCGCGCGTCTTCG---GGGTCTTGCGCCTTGGGCTC  943
           |||||||||||||||||  |||||   ||||||||| ||||  ||||||| ||||||||||
Db    621  CTACATGCTATTGGCCACGGGGCATCTTCCAGCGTCGTTGGGCTGTCGTCTGTGTGTCTGGCTC  680

Qy    944  CACTCTAGGCGTGCCGCGCTTCTGCCGCCGCTCCTCGCGCCTCACCACTCAGTA  1003
           ||||||||||||||| |||  |||||||||||||||||||| ||||| |||||||
Db    681  CACATCACTCACCTGTCCCCTGGACCTGGGGCGTTCTGTTGTGCGTGCTACTCACCA  740

Qy   1004  CCGCCGGCCGGCCCTTCTGGGCGTCGGTCAACGGCGGTCCTGGTCGTTCTTCCTCGGAGGGGC  1063
           ||||||||||||||| |||||||||||  |||  ||||||||| ||||    ||||||||||
Db    741  TCGGGCCCTGCCTTCTGGGGATCACATTGACCACCAGAGGACTGCTGTTG  800

Qy   1064  CGTGGTGAGTCTCCAGTATGTGTTCGGGCCAAGCCGTACCCTTCTGGACCAAAGGC  1123
           ||||||||| ||||||||||||||||   ||||  |||||||||| ||||||||||
Db    801  TATGGCGGTGGGCCCACAGGATGCAGCCTCAGAAGGCTTTCTTCTTCAACCAGAGTGT  860

Qy   1124  CAAGGA  1129
           ||||||
Db    861  GGATGA  866
```

```
Qy    323  CTTCAGCGTTCCACTGCTCATCGTTATTCTAGTGTTGTTTGGGCTCTTAGCACAAGCTTCCT  382
           |||||||||||||||  |||  |||||  ||     | |||| | |||||| || |||||||
Db    126  GATGGACACCACTTTGGCCAGCATCATCAGTGATCTTTTCTGACTGCACTGGGTTCTTCAT  185

Qy    383  GCTCATCTTGCCGGGGGATCCGTTGGCCACTGCGGCTGGTGGTGTTCTT  442
           || |||||||||||||||||||||||||||||   ||||||||||||||
Db    186  CGTCATCCTGCCCTGGCATTCGGGGAAAGAGGAGGCTGTTGGCTTCGGGTGTGGAC  245

Qy    443  CAGTCTGTTCATAGGCGCAGAAAATTGTGGCTGTGTGCACTTCAGTGCGAGAATGGTTCGTGGG  502
           |||||||||||||||  |||  |||||||||||||  || |||||||||||| |||||||||||
Db    246  CAGCTTATTCATCGGGGCTGCAATCCTGGCTGTGAATTTCAGTTCTGAGTGGTCTGTTGGG  305

Qy    503  TACAGTGAACACCAACACATCCTAACAAAGCCTTCAGCGCGGCGGAGCGGGGGCCCGGTGT  562
           |||||||||||| |||||  ||||||| |||||||||||||||||||||| |||||||||
Db    306  CCAGGTCAGCACCAACACATCACAAGCCCTTCAGTTCTGCGGGATCGGGTTACAGCCCGGTGT  365

Qy    563  CGGTCGTGCTGGTGGGGCCTGGAGGGGCATTAATATTACACTCACAGGGACCCCAGTGCATCA  622
           |||||||||||||||||||| ||||||| |||||||||||| || |||||||||||| |||
Db    366  TGGGCTGCAGGTCGGGTGTGGGCTGGGAGTCAACATCACACTCACAGGGAACCCCGTGCAGCA  425

Qy    623  GCTGAACGAGACCAATTGACTACAACAGGAGTTCAACGAGAGTTCAACATCACACTGGGTGAAAGAGAATTACGC  682
           ||||||  ||||||  |||||||| ||| ||||| || |||||||||||||||||||| ||  ||| || |||||
Db    426  GCTGAATGAGACCAATCAATTACAACGAGAGTTCACCTGGCGTCACAGGGGAACTATGC  485

Qy    683  CGCGGAGTACGCGAACGCACTGGAGAAGGGGCTGCCGGACCCAGTGCTTCTACCTGGGCGGA  742
           ||||||| ||   ||| |||||||||||||||||||||||||||| |||| |||| ||||
Db    486  TGAGGAGTGTGCAAAGGCTCTCGGAGAAGGGGCCGGACCCGTGTCTTACCTTAGCTGA  545

Qy    743  GAAGTTCACACGAGCTAGCCCTTGCGGCCTGTACCACCAGTACCACCTGGCGGGGGGGCACTA  802
           ||||||||| |||||||  ||||| |||| |||||||||||| |||||||| ||||||| ||
Db    546  GAAGTTCACTCAGAAGCCCATGTGGCCTATACCGGCCAGTACCGGCTGGGCGGGCACTA  605

Qy    803  CGCCTCGGCGCACCGGCTATGGGGGGTTCTGCTTCTGGCTGCTTCTCCAAGCTGTGTCTC  862
           ||||| ||| ||||| ||||| ||||||||| |||||||||||   |||| |||||||||
Db    606  CACCTCAGCCATGCTATGGGTGGCCAATTCCTCTGCTGCGTCGGCGGCAAATGATGCTCTC  665

Qy    863  CACGCCGGCGCCCGGCTCTACGGAGGGCCTGCACTGCTGACCACCGGGAGCCCTTCGCGCGCTCTT  922
           |||||||||| || ||| || |||||| ||| || ||||||||| |||||||| ||| ||
Db    606  GAAGTCACTCCAAGAAGCCCATGTGGCCTACATGCTATTGGCCACGGGGCATCTTCCAGCGTGTT  725

Qy    923  CG---GGGTCTTCGGCCTTGGGCTCCCATCTCTAGGGTGCGGCCTCCGCCGGCTCCGCCTAGG  979
           ||   |||||||||||||| |||||||||  |||||||| |||||||||| ||||||||
Db    726  GGCTGCGTGCTCTTTCTTCTCCATGGCCACCACTCAACTACCCCGTCCCTGCACCGGGT  785

Qy    980  CTCCTCCCGGCGTCACCACTCAGATGACCACGGGGCCGCCCCTTCGGGGTCACGCTGGGCAACCGGGGT  1039
           |||||| | |||||||||||  ||||| ||| ||||||||||||| ||||||||||||| |||||||
Db    786  CGCTTCTGTGCTGCATACTCACCATGGGGCCTGCCTTCTTCTTCAACATTGACCACACAGGGACT  845

Qy   1040  CCTGTGCCTCTTCCTCGGAGGGGGCCGTGGTGTGAGTCTCCAGTATGTGTTCGGGCCAAGCGCGCTCT  1099
           ||||||||| |||||||| || |||||||||||||| || || |||||||| || || |||||||
Db    846  GCTGTGTGTGCTGCTGTGGGCGGTGGGCCCACAGGATGCAGCCTCAGAAGGCTTTCTTCTTCAACAGGCT  905

Qy   1100  TCGGCACCCTTCTGGACCAAAGGCCAAGGA  1129
           |||||||||| |||||||||||||||||||
Db    906  GAAGGCTTTCTTCAACCAGAGTGTGGATGA  935
```

RESULT 2
US-09-578-063-53
; Sequence 53, Application US/09578063
; Patent No. 6764677
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A
; APPLICANT: Barnes, Thomas M
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 210147.0023/6U1
; CURRENT APPLICATION NUMBER: US/09/578,063
; CURRENT FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 2133
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-578-063-53

Query Match   25.0%;   Score 354.8;   DB 3;   Length 2133;
Best Local Similarity   63.8%;   Pred. No. 9.3e-78;
Matches 555;   Conservative   0;   Mismatches 312;   Indels   3;   Gaps   1;

RESULT 3
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA

```
Qy    263  CAGCATGACCCTGTGGGAACGGGCGGTACTGCCTGTTTTACCCCCAGCCCCCGGCGCATGCGCGCAGG  322
           ||||||| ||| |||| | ||||||||| |||  || |||| | ||||||||||||||||||||
Db     66  CAAGATGGCTACTTTGGGGACACACATTCCCCTTTCTATGCTGTGGCCCCAAGCCAACCTTCC  125
```

; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; PRIOR APPLICATION DATA:
; CLASSIFICATION: 435
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pT2gpt-Fls
US-08-232-463-14

Query Match   4.3%;   Score 61; DB 2; Length 7218;
Best Local Similarity   1.9%; Pred. No. 3.5e-05;
Matches 7; Conservative 221; Mismatches 131; Indels 0; Gaps 0;

Qy   761 CCCTTGCGCGCCTGTACCACCAGTACCACCTGGGGCGGGGACACTACGGCTCGGCCACCGCTATG 820
          :::::: :::: :::: ::: :::::: :::::::::::::::::::::::::::::::::::::: :
Db   1084 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1143

Qy   821 GGTGGCGTTCTGCTTCTGGCTCCTCCAAGTGTGCTGCTCTCCACGCCGCGCCCCGCTCTA 880
          : ::::::: ::::::::::::: :::::::::::::::::::::::::::::::::::::::
Db   1144 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1203

Qy   881 CGGAGGCCTCGGCCACTGCCTGACCACCGGGAGCCCTTCGGGCTCCTTCGGGGGTCTTCGCCTTGGC 940
          :::::: ::::::: :::::::::::::::::::::::::: ::::::::::::::: :
Db   1204 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1263

Qy   941 CTCCATCTCTAGCGTGCCGGCTCTGCCCGCTCCTCCGGCTGCCTCTTCCTCCGGAGG 1000
          ::::: :::::: :::::::::::::::::::::::::::::::::::::: :: :::
Db   1264 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1323

Qy   1001 GTACGGCGCGCGCCTTCTGGGTCACGCGTGGCAACCGGCGTCCTGGCTCCTCCTCGGAGG 1060
          : : ::: ::
Db   1324 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1383

Qy   1061 GGCCGTGGTGAGTCTCCAGTATGTTCGGCCCAGGCGGCTTCTTCGGCACCCTTCTGGACCAAA 1119
          ::::: :::: ::: :::::::::::::::::::::::::::::::::: ||||||
Db   1384 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYVGTACCAAA 1442

RESULT 4
US-09-252-991A-1304/c
; Sequence 1304, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1304
; LENGTH: 1098
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1304

Query Match   3.5%;   Score 50.2; DB 3; Length 1098;
Best Local Similarity 50.6%; Pred. No. 0.01;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy   838 GGCTCCTCTCCAACGTGCTGCTCTCCACGCCGCGCCCCGCTCTACCGGAGGCCTCGGCCACTGC 897
          ||||| ||| ||| ||||||| || ||||||||||||||| ||||||| ||| | |||
Db   602 GACTCCTGCTCAATCGGCTCGCGGGCGCGCGCGCCGGAGCTGGTCTCGGGGGCGGTGATCGGTGC 543

Qy   898 TGACCACCGGAGCCCTTCGCGGCTCTTCGGGGGTCTTCGGGCTCCATCTCTAGCGTGC 957
          ||||| ||| ||||| ||||| |||||| || ||| || | ||| ||
Db   542 TGGCGCGCCGGCGACTCGGGCCCGGAAGCGGCAGACCCTCCCCGGCGCGCGCG 483

Qy   958 CGCTCTGCCCGCTCCCGCTAGGCTCCTCCGGCGGCTCACCACTCAGTAGCGCGCTTCT 1017
          ||||| || |||| ||||||||||| ||||| | |||||| ||||||| ||| |||
Db   482 TGCTCGCCCGGCCGGTGTGCGCCGGAAGCACCCGGCGCGCACCCGCGCGCGAGGCCT 423

Qy   1018 GGGTCACGCTGGCAACCGGCGGGTCCTGTGCCTCCTTGGCCAGCCGGCGCCTTCT 1076
          | ||||||||||||| || |||| | |||||| ||| || || | |||
Db   482 TGCGCTGGCCGCGCACGGGAGGGCCTCGGCGGATGCCCTATGGCACCCTGCCGCGGGGGGGTC 364

RESULT 5
US-09-252-991A-1255
; Sequence 1255, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1255
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1255

Query Match   3.5%;   Score 50.2; DB 3; Length 1257;
Best Local Similarity 50.6%; Pred. No. 0.01;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy   838 GGCTCCTCTCCAACGTGCTGCTCTCCACGCCGCGCCCCGCTCTACCGGAGGCCTCGGCCACTGC 897
          ||||| ||| ||| ||||||| || ||||||||||||||| ||||||| ||| | |||
Db   333 GACTCCTGCTCAATGCGCTGCGCGCGCGCGCCGGAGCTGGTCTCGGGGGCGGCGATGGGTGC 392

Qy   898 TGACCACCGGAGCCTTCGGCGCTCTTCGGGGTCTTTCGCCTTGGCCTTCAATCTCTAGCGTGC 957
          | ||||| ||| ||||| ||||| |||| || ||||| || ||| ||
Db   393 TGGCGGCCGCGGCTCGGGGCCGGAACGCACGGGACCCTCGCCGGCGGCGCACCGGCGCGCG 452

Qy   958 CGCTCTGCCCGCTCCCGCTAGGCTCTTCACCACTCAGTACGGCGCGCCTTCT 1017
          ||||| || |||| ||| || || | |||||| ||||| ||| | | |||
Db   453 TGCTCGCGCCCGGCTTGTTCCGGAGAACACCCGCGCGCGCAGCCGCGCGAGGCCT 512

Qy   1018 GGGTCACGCTGGCAACCGGCGTCCTGTGCCTCTTCCTCCGGAGGGGGGGGTGGGTGAGTCTC 1076
          | ||||||||||||| || |||| | ||||| ||| | ||||| || || ||
Db   1018 GGGTCAGGCTGGCGAACCGGCCGTCCGGGGTGGGTGGGTGAGTCTC 1076

```
Db   513  TGCGCCTGGCCGAGCGGACCGGCCTCGGGCGCATTCGCCTATGGCACCCTGCCGGGGGTC  571
```

RESULT 6
US-09-252-991A-1163
; Sequence 1163, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1163
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa

Query Match   3.5%; Score 50.2; DB 3; Length 1746;
Best Local Similarity   50.6%; Pred. No. 0.011;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

```
Qy    838  GGCTCCTCTCCAACGTGCTGCTCTCCACGCCGGCCCCGCTCTACGGAGGCCTGGCACTGC   897
           ||| |||| || |  | | || | ||| |  |  ||| ||| ||||| ||||||||||||
Db   1283  GACTCCTGCTCAANTGCGCTGCGCGGCGGTGCCCGGAGCTGGTTGCTCTGGGCGGCTGANG 1342

Qy    898  TGACCACCGGAGCCTTCGCCGCTCTTCGGCGGTCTTCGCTGGCTCCATCTCTGGCCGTGC   957
           || | |||| |||| ||| || |  |  |||  ||| || ||  |  |  |  | ||||||
Db   1343  TGGCGGCCGGGGCTCGGGCGCCGAACCCGCGGCGACCCTCGCGGCGGGGCGGGGCGGGCG  1402

Qy    958  CGCTCTGCCCGCGTCCGCCCTAGGCTTCACCACTCAGTACGGCGCGCCGCCTTCT       1017
           ||| | || |  | | ||| | || ||| ||| |  |  ||| || ||||| |
Db   1403  TGCTCGGCGGCGGCGTTCGCCGAAGACCCTGAGAACACCCCGCGGCAGCCGCCGGAGGCCT 1462

Qy   1018  GGGTCACGCTGGCAACCGGCGGGTCCTTCGTGTGCCTCTTCCTCGGAGGGGGCCCGGTG   1076
           |||||| | |||||||||| || |  ||  | | ||| |  |  || |  |   |||||
Db   1463  TGCGGCTGGCCGCGCGCAGCGGACGCCTCGCGGGCCATTCGCCCTATGGCACCCTGCCGCGGGGGTC 1521
```

Query Match   3.4%; Score 48.2; DB 3; Length 1746;

RESULT 7
US-09-252-991A-1350/c
; Sequence 1350, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1350
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa

Query Match   3.4%; Score 48.2; DB 3; Length 396;
Best Local Similarity   52.2%; Pred. No. 0.024;
Matches 107; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

```
Qy    838  GGCTCCTCTCCAACGTGCTGCTCTCCAGCGCCGGCCCCGCTCTACGGAGGCCTGGCACTGC  897
           ...
Db    163  ...                                                           ...
```

RESULT 8
US-09-614-912-35
; Sequence 35, Application US/09614912
; Patent No. 6677502
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Caimi, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes
; FILE REFERENCE: BB1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,401
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/146,650
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/170,906
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: 60/172,959
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/172,946
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 35
; LENGTH: 1436
; TYPE: DNA
; ORGANISM: Zea mays

Query Match   3.4%; Score 48; DB 3; Length 1436;
Best Local Similarity   46.8%; Pred. No. 0.038;
Matches 187; Conservative 0; Mismatches 210; Indels 3; Gaps 1;

```
Db    217  GACTCCTGCTCAANTGCGGCTGCGCGGCGGTGCCGGGGCGGGGTGCCGGAGCTGGGTCTTGCGGCGGTTGC  158
Db    157  TGGCGGCCGGGGCTCGGGCCGAACGCGGGACCCTCGCGGCGGGGCGGGGCGGGCGCACCACGGGCGGCG    98
Db     97  TGCTCGGCGGCGGCGTTCGCCGAAGACCCTGAGAACACCCCGCGGCAGCCGCCGGAGGCCT           38
Db     37  TGCGGCTGGCCGCCCGCCAGCCGGACGCGT                                          13
```

```
Qy    549  GTTACAGCCCGTGTGTCGGGTCTGCTCCGTGGAGGGCCTGGAGGGCATTAATATTACACTCACAGGG  608
Db     43  GCTTCCTCCCTTCGCCGCTCCTCACCATGGTGCTGGCTGGCGACGATCCCCTTCCCTTGTG        102

Qy    609  ACCCCAGTGCATCAGGTGAACGGAGACCATTGACTACAACGACAGTTCACCTGGCGTCTG  668
Db     43  GCTTCCTCCCT...                                              102

Qy    669  AAAGAGAATTACGCGCGGCGGAGTACGGCGAACGCACTGGAGAAGAGGCGTGCCGGACAGTG  728
Db    163  GACGAGAGCCCCACCCGGCGGGAGAAGTTCACACCG---AGTAGCCCTTCGCGGCTGTACCACCAGTAC  222

Qy    729  CTCTACCTGGGCGGAGAAAGTTCACACCG---AGTAGCCCTTCGCGGCTGTACCACCAGTAC 785
```

RESULT 9
US-09-252-991A-11029/c
Sequence 11029, Application US/09252991A
Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11029
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11029

Query Match   3.3%;   Score 47.6; DB 3; Length 420;
Best Local Similarity   48.2%; Pred. No. 0.035;
Matches 134; Conservative   0; Mismatches 144; Indels   0; Gaps   0;

Qy   793 CGGGACACTACGCCTCGGCCACGCTTATGGGTGGCGGTTCTGCTTCTGGCTCCTCCAACG 852
          |||| | ||| || | || ||||||| | || |||| | || | | ||||||
Db   382 CGGTGGAAAACGCTTCCTTCATGCCCTGGCTGGTCGGCACTGACCACCGGAGCCT 323

Qy   853 TGCTGCTCTCCACGCGCCCCGCTCTACGGAGGCCTGGCACTGACGATCGGCTGG 912
          ||||||||| || |||| |||| | ||| ||||||||||||||| ||||||||
Db   322 CGGTGACCGAGAAGCGCGCCGTGTTCAAGAGCTGACACCGGTCGTCTTGGGCATCGGCGTT 263

Qy   913 TCGCGGCTCTTCGGGGGTCTTCGCGCTCCCATCTCTAGCGTGCCGCTCC 972
          | || | ||| |||||| | | | ||||||| | || |||||||||||
Db   262 TCTCCCTGAGCCTGCTGGGCACCTTCCTGGTCGGCTCACCTCGGTGCACG 203

Qy   973 GCCTAGGCTTCCTCCGCGGCTCACCACTCAGTACGCGGCGGCGCCTTCTCGGTCAGCGCTGGCAA 1032
          | |||||| ||||| ||| ||| | | || || | | | || ||| || ||||| |||||
Db   262 CCTTCGCGCTGCTGGGACACGCGGGGGTGTTCATCCTCGCCTTCCTGCTCGGTGGTCG 203

Qy   1033 CCGGGCGGTCCTCGTGCCTCTTCCTCGGAGGGGGCCGGTGGGTG 1070
          ||| | || | ||| | || | || | | | ||| |||||
Db   202 CCTTCGCGCTCCGACCCGACGCGGGGGGTGTTCATCCTCGCTGCTGGTGGTCG 143

RESULT 10
US-09-252-991A-11226
Sequence 11226, Application US/09252991A
Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11226
; LENGTH: 1683
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11226

Query Match   3.3%;   Score 47.6; DB 3; Length 1683;
Best Local Similarity   48.2%; Pred. No. 0.049;
Matches 134; Conservative   0; Mismatches 144; Indels   0; Gaps   0;

Qy   793 CGGGACACTACGCCTCGGCCACGCTTATGGGTGGCGGTTCTGCTTCTGGCTCCTCCAACG 852
          |||| | ||| || | || ||||||| | || |||| | || | | ||||||
Db   864 CGGTGGAAAACGCTTCCTTCATGCCCTGGCTGGTCGGCACTGACCACCGGAGCCT 923

Qy   853 TGCTGCTCTCCACGCGCCCCGCTCTACGGAGGCCTGGCACTGACGATCGGCTGG 912
          ||||||||| || |||| |||| | ||| ||||||||||||||| ||||||||
Db   924 CGGTGACCGAGAAGCGCGCCGTGTTCAAGAGCTGACACCGGTCGTCTTGGGCATCGGCGTT 983

Qy   913 TCGCGGCTCTTCGGGGGTCTTCGCGCTCCCATCTCTAGCGTGCCGCTCC 972
          | || | ||| |||||| | | | ||||||| | || |||||||||||
Db   984 TCTCCCTGAGCCTGCTGGGCACCTTCCTGGTCGGCTCACCTCGGTGCACG 1043

Qy   973 GCCTAGGCTTCCTCCGCGGCTCACCACTCAGTACGCGGCGGCGCCTTCTCGGTCAGCGCTGGCAA 1032
          | |||||| ||||| ||| ||| | | || || | | | || ||| || ||||| |||||
Db   1044 CCTTCGCGCTGCTGGGACACGCGGGGGTGTTCATCCTCGCCTTCCTGCTCGGTGGTCG 1103

Qy   1033 CCGGGCGGTCCTCGTGCCTCTTCCTCGGAGGGGGCCGGTGGGTG 1070
          ||| | || | ||| | || | || | | | ||| |||||
Db   1104 GCGGGCTCGCTGACCCTGTTCGCCCGTGCGCGGCGCCGGTG 1141

RESULT 11
US-09-252-991A-11111
Sequence 11111, Application US/09252991A
Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11111
; LENGTH: 1974
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11111

Query Match   3.3%;   Score 47.6; DB 3; Length 1974;
Best Local Similarity   48.2%; Pred. No. 0.051;
Matches 134; Conservative   0; Mismatches 144; Indels   0; Gaps   0;

Qy   793 CGGGACACTACGCCTCGGCCACGCTTATGGGTGGCGGTTCTGCTTCTGGCTCCTCCAACG 852
          |||| | ||| || | || ||||||| | || |||| | || | | ||||||
Db   731 CGGTGGAAAACGCTTCCTTCATGCCCTGGCTGGTCGGCACTGACCACCGGAGCCT 790

Qy   853 TGCTGCTCTCCACGCGCCCCGCTCTACGGAGGCCTGGCACTGACGATCGGCTGG 912
          ||||||||| || |||| |||| | ||| ||||||||||||||| ||||||||
Db   791 CGGTGACCGAGAAGCGCGCCGTGTTCAAGAGCTGACACCGGTCGTCTTGGGCATCGGCGTT 850

Qy   913 TCGCGGCTCTTCGGGGGTCTTCGCGCTCCCATCTCTAGCGTGCCGCTCC 972
          | || | ||| |||||| | | | ||||||| | || |||||||||||
Db   851 TCTCCCTGAGCCTGCTGGGCACCTTCCTGGTCGGCTCACCTCGGTGCACG 910

```
Qy    973 GCCTAGGGCTTCCTCCGCGCTCACCACTCAGTACGGCGCGCGCCCTTCTGGGTCACGCTGGCAA 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    911 CCTTCGCCCTGACCCGAGCGCGGGGGTGTTCATCCTCGGCCTTCCTGCTGCTGGTGGTCG 970

Qy   1033 CCGGGCGTCCTGTGCCTCTTCCTCTCGGAGGGGCCGTGGTG 1070
              ||||||||
Db    971 GCGGGCTCGCTGACCTGTTTCGCCCTGCCTGCGCGCGCGGGTG 1008
```

RESULT 12
US-09-252-991A-10956/c
; Sequence 10956, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;   APPLICANT: Marc J. Rubenfield et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;   TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 107196.136
;   CURRENT APPLICATION NUMBER: US/09/252,991A
;   CURRENT FILING DATE: 1999-02-18
;   PRIOR APPLICATION NUMBER: US 60/074,788
;   PRIOR FILING DATE: 1998-02-18
;   PRIOR APPLICATION NUMBER: US 60/094,190
;   PRIOR FILING DATE: 1998-07-27
;   NUMBER OF SEQ ID NOS: 33142
;   SEQ ID NO 10956
;   LENGTH: 2283
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10956

Query Match   3.3%; Score 47.6; DB 3; Length 2283;
Best Local Similarity   48.2%; Conservative 0; Mismatches 0.053; Indels 0; Gaps 0;
Matches 134;

```
Qy    793 CGGGACACTACGGCCTCGGCCCACCGCTATGGGTGGCGTTCGTTCTTGGTCTTCCTCCAACG 852
              || ||          |||||||||||||||||||| |||||||||||||||||||||||||
Db   2081 CCGGTGGAAAACGGTTCCTTCATGCCCTGGCTGGTCGGCGTTCATCCTCGGCCTTCCTTGTG 2022

Qy    853 TGCTGCTCTCCACGCGCGGCCCCCGCTTCTACGGAGGCCTGGCACTGCTTGACCACCGGAGCCT 912
              |||||||||||| |||||||||||||||| |||||||||||  ||||||||||||||||
Db   2021 CGGTGACCGAGAAGCGCGCGGCGTTCAAGAGCTGGACACCGTGCTTCTTCTGGCGATCGCCGCTT 1962

Qy    913 TCGGCGCTCTTCTGGGGGTCTTCGCCTTGGCCTTCCATCTCTTAGGCGTGCCGGCTTCTCCGCGCTCC 972
              |||||||||||||||||||||||||||||| ||||||||| |||||||||||||||||||
Db   1961 TCTCCCCTGAGCCTGCTGGGCACCTTCCTGGTCGCCGTGCTCCACCTCGGTCGACG 1902

Qy    973 GCCTAGGCGTTCCTCCGCGGTCACCACTCAGTACGGCGCGCCGTCAGGCTGGCAA 1032
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   1961 CCTTCGCCCTCCGACCCGGAGCGCGGGGGTGTTCATCCTCGGCCTTCCTGCTGCTGGTGGTCG 1842

Qy   1033 CCGGGCGTCCTGTGCCTCTTCCTCTCGGAGGGGCCGTGGTG 1070
              |||||||
Db   1841 GCGGGCTCGCTGACCCTGTTTCGCCCTGCGCGCGCGGGTG 1804
```

RESULT 13
US-09-614-912-49
; Sequence 49, Application US/09614912
; Patent No. 6677502
; GENERAL INFORMATION:
;   APPLICANT: Allen, Steve
;   APPLICANT: Rafalski, Antoni
;   APPLICANT: Orozco, Buddy
;   APPLICANT: Miao, Gou-Hau
;   APPLICANT: Famodu, Omolayo O.
;   APPLICANT: Lee, Jian Ming
;   APPLICANT: Sakai, Hajime
;   APPLICANT: Weng, Zude
;   APPLICANT: Caimi, Perry G
;   APPLICANT: Anderson, Shawn
;   TITLE OF INVENTION: Plant Metabolism Genes

; FILE REFERENCE: BB1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,401
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/146,650
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/170,906
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: 60/172,959
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/172,946
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 49
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1309)
; NAME/KEY: unsure
; LOCATION: (1339)
; NAME/KEY: unsure
; LOCATION: (1351)
; NAME/KEY: unsure
; LOCATION: (1402)
; NAME/KEY: unsure
; LOCATION: (1429)
US-09-614-912-49

Query Match   3.3%; Score 46.4; DB 3; Length 1431;
Best Local Similarity   46.5%; Conservative 0; Mismatches 0.094; Indels 3; Gaps 1;
Matches 186;

```
Qy    549 GTTACAGCCCGTGTTCGGTCGTCCTCGTCGGCGTCGGGCGGGGGCCATTAATTATTACACTCACAGG 608
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db     33 GCTTCCTCCCTCCTCCCTCCACCATGGTGCTGGCGCCACGGATCCCCCTCCCCTTGTG 92

Qy    609 ACCCCAGTGCATCAGCTGAAACGGAGGACCAATTGACTACAACGAGACGAGAGTTCACCTGGCTG 668
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db     93 TTCGACGCTGCCCCGCCTGAGCGGCGCCTTCCCGACAATCCCGCAGCAGTTCATCTCGGCCGGCGGG 152

Qy    669 AAAGAGAATTACCGCGCGGCGTACGCAGTACGGAGAAGGCGGAGGACCCCAGTG 728
              ||||||||||||||||||||| |||||||||||||||||||||||||| ||||
Db    153 GACGAGAGCCCCACCCCGGACGTCCGCCGGAGGAGTGGCCGGTGCCGCTCATCACGACCTCTCC 212

Qy    729 CTCTACCTGGCGGGAGGAAGTTCACACCG----AGTAGCCCTTGCCGGCGTGTACCACCAGTAC 785
              |||||||||||||||||| ||||| |||    ||||||||||||||||||||||
Db    213 GGGGAAGCCGCCGAGGTGGTCCGGCGACGGTCGGCGCGCCCTGCGACCTGGCGGCTTCTTC 272

Qy    786 CACCTGGCGCGGGACACTACGCCCTCGGCCACAGCTATGGGTGGCGTTCGTTCGCTTCTC 845
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 CAGGTGGTGGGGACGGGCATGGACGGGCGGCGGGGGGCTGACGCGCGAGGCCCCACCGGTTCGCATGGAC 332

Qy    846 TCCAACGTGCTGCTCTCCACGCCCGGCGTCTACGGAGGCGTTGGCACTGCTGACCACC 905
              ||||||||||||||||||||||||||||||| ||||||||
Db    333 GCTTCTTCACCGTGCTGCCGCTCCCGGACAAGCAGCAGCGCCGGCCGCGGCGCGGGGACAGC 392

Qy    906 GGAGCCTTCGCGCTCTTCGGGGTCTTCGCCTTGGCCTCCA 945
              |||||||| |||| ||||||||||||||||||
Db    393 TGCGGCTACGCCCAGCCAGCTTCACGGGGCCCGGTTCGCGGTCCA 432
```

RESULT 14
US-09-949-016-88511
; Sequence 88511, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```
;   TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;   TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;   FILE REFERENCE: CL001307
;   CURRENT APPLICATION NUMBER: US/09/949,016
;   CURRENT FILING DATE:  2000-04-14
;   PRIOR APPLICATION NUMBER: 60/241,755
;   PRIOR FILING DATE:  2000-10-20
;   PRIOR APPLICATION NUMBER: 60/237,768
;   PRIOR FILING DATE:  2000-10-03
;   PRIOR APPLICATION NUMBER: 60/231,498
;   PRIOR FILING DATE:  2000-09-08
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   NUMBER OF SEQ ID NOS: 207012
;   SEQ ID NO 88511
;     LENGTH: 601
;     TYPE: DNA
;     ORGANISM: Human
US-09-949-016-88511

Query Match             3.2%;  Score 46;  DB 3;  Length 601;
Best Local Similarity   57.7%;  Pred. No. 0.095;
Matches   82;  Conservative   0;  Mismatches  60;  Indels   0;  Gaps   0;

Qy   1035 GGGCTCCGTGCCTCTTCCTTCGGAGGGGCCGTGGTGAGTCTCCAGTATGTTCGGCCCAGC 1094
               | | | | | ||||| ||| | |||| |  ||| | | ||| |  || || |||| |||
Db     63 GGGGTGTTGATCCTGGGAGTAGGGGAGACAGAGTTGGGGCCCTAGGATGCCTGGCACAGC 122

Qy   1095 GCTCTTCGGCACCCTTCTTCGGACAAAAGCGCCCAAGGACTCGAGCCAGCAGAGAGGGCTCA 1154
              ||||  | |||||  || | ||||  ||||||||  |||| | ||| ||||||||  |||
Db    123 AGGGTCCTCACCCTGTGGGCCCAATGTGATAAGCACTGAGGCCGGAGATAAGGGGCTTT 182

Qy   1155 CCTCTTATCCTCGGCGACCCAC 1176
              || ||||||||| ||| ||| ||
Db    183 AGAGTAGGCCGCTGGGGAGCCTC 204


RESULT 15
US-09-949-016-14248/c
;   Sequence 14248, Application US/09949016
;   Patent No. 6812339
;   GENERAL INFORMATION:
;   APPLICANT: VENTER, J. Craig et al.
;   TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;   TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;   FILE REFERENCE: CL001307
;   CURRENT APPLICATION NUMBER: US/09/949,016
;   CURRENT FILING DATE:  2000-04-14
;   PRIOR APPLICATION NUMBER: 60/241,755
;   PRIOR FILING DATE:  2000-10-20
;   PRIOR APPLICATION NUMBER: 60/237,768
;   PRIOR FILING DATE:  2000-10-03
;   PRIOR APPLICATION NUMBER: 60/231,498
;   PRIOR FILING DATE:  2000-09-08
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   NUMBER OF SEQ ID NOS: 207012
;   SEQ ID NO 14248
;     LENGTH: 8991
;     TYPE: DNA
;     ORGANISM: Human
US-09-949-016-14248

Query Match             3.2%;  Score 46;  DB 3;  Length 8991;
Best Local Similarity   57.7%;  Pred. No. 0.19;
Matches   82;  Conservative   0;  Mismatches  60;  Indels   0;  Gaps   0;

Qy   1035 GGGCGTCCTGTGCCTCCTTCCTTCGGAGGGGGCCGTGGTGAGTCTCCAGTATGTTCGGCCCAGC 1094
               | | | | | ||| ||| | ||| |  ||| | | ||| |  || || |||| |||
Db   3532 GGGGTGTTGATCCTGGGAGTAGGGGAGACAGAGTTGGGGCCCTAGGATGCCTGGCACAGC 3473

Qy   1095 GCTCTTCGCACCCTTCTTCTGGACCAAAAGCGCCAAGGACTGCAGCCAGCAGAGAGGGGCCTCA 1154
              ||||  | |||||  || | ||| ||||||| | |||| | ||| ||||||||  |||
Db   3472 AGGGTCCTCACCCTGTGGGCCCAATGTGATAAGCACTGAGGCCGGAGATAAGGGGCTTT 3413
```

```
                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          January 21, 2006, 22:28:47 ; Search time 1286 seconds
                 (without alignments)
                 9137.470 Million cell updates/sec

Title:           1421
Perfect score:   1421
Sequence:        1 gaaccaggaaagtaaccggct.........cttcttcttcttgtcttct 1421
Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched:        9793542 seqs, 4134689005 residues
Total number of hits satisfying chosen parameters:    19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA_Main:*
                 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
                 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
                 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
                 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
                 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
                 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
                 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
                 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1421	100.0	1421	5	US-10-187-657-2	Sequence 2, Appli
2	1421	100.0	1450	8	US-10-643-795A-5	Sequence 5, Appli
3	1421	100.0	1450	9	US-10-948-518-5	Sequence 5, Appli
4	1418.6	99.8	1474	6	US-10-264-237-1097	Sequence 1097, Ap
5	1415.8	99.8	1418	5	US-10-187-657-4	Sequence 4, Appli
6	1415.2	99.6	1420	5	US-10-773-236-187	Sequence 187, App
7	1412	99.4	1491	9	US-10-187-657-9	Sequence 9, Appli
8	693	48.8	1594	5	US-10-187-657-5	Sequence 5, Appli
9	459.6	32.3	522	5	US-10-773-236-69	Sequence 69, Appl
10	413	29.1	580	9	US-10-450-763-2367	Sequence 2367, Ap
11	402.0	28.3	406	9	US-09-925-065A-72673	Sequence 72673, A
12	396.6	27.9	1376	6	US-10-094-749-1195	Sequence 1195, Ap
13	391	27.5	2684	6	US-10-187-657-7	Sequence 7, Appli
14	385.8	27.1	450	6	US-09-759-130B-424	Sequence 424, App
15	354.8	25.0	1029	3	US-10-042-431-54	Sequence 54, Appl
16	354.8	25.0	1029	7	US-10-741-790-424	Sequence 424, App
17	354.8	25.0	2133	3	US-09-759-130B-423	Sequence 423, App
18	354.8	25.0	2133	7	US-10-042-431-53	Sequence 53, Appl
19	354.8	25.0	2133	5	US-10-741-790-423	Sequence 423, App
20	322	22.7	506	5	US-10-187-657-6	Sequence 6, Appli
21	300	21.1	346	6	US-10-187-657-3	Sequence 3, Appli
22	245.8	17.3	2684	6	US-10-094-749-1195	Sequence 1195, Ap
23	217.4	15.3	395	4	US-09-925-065A-100165	Sequence 100165,
24	217.4	15.3	395	4	US-09-925-065A-145294	Sequence 145294,
25	200.6	14.1	439	6	US-10-027-632-91970	Sequence 91970, A
26	200.6	14.1	439	6	US-10-027-632-91970	Sequence 91970, A
27	199.4	14.0	232	3	US-09-783-590-9492	Sequence 9492, Ap
28	190	13.4	277	4	US-09-925-065A-100166	Sequence 100166,
29	170	12.0	250	5	US-10-187-657-8	Sequence 8, Appli
30	161.8	11.4	1338	9	US-10-450-763-3511	Sequence 3511, A
31	65.6	4.6	1602	10	US-11-097-143-29060	Sequence 29060, A
32	65.6	4.6	1808	10	US-11-097-143-8903	Sequence 8903, Ap
33	64	4.6	512	5	US-10-437-963-340	Sequence 340, App
34	55.2	3.9	1783	7	US-10-767-701-13710	Sequence 13710, A
35	54.8	3.9	555	7	US-10-437-963-10408	Sequence 10408, A
36	53	3.7	3024	7	US-10-437-963-35623	Sequence 35623, A
37	52.6	3.7	1087	8	US-10-425-115-140838	Sequence 140838,
38	52	3.7	8134	7	US-10-437-963-93461	Sequence 93461, A
39	51.6	3.6	1280	7	US-10-437-963-40658	Sequence 40658, A
40	51.6	3.6	1506	7	US-10-437-963-40658	Sequence 40658, A
41	51.2	3.6	1123	7	US-10-425-114-15938	Sequence 15938, A
42	51.2	3.6	1454	8	US-10-425-115-46357	Sequence 46357, A
43	51.2	3.6	2508	7	US-10-437-963-2740	Sequence 2740, Ap
44	50.8	3.6	1455	6	US-10-156-761-5889	Sequence 5889, Ap
45	50.6	3.6	9025608	6	US-10-156-761-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-10-187-657-2
; Sequence 2, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1421
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 4901066CB1
US-10-187-657-2

Query Match             100.0%;  Score 1421;  DB 5;  Length 1421;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1421;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1 GAACCAGGAAAGTAACCGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGCTAGAAAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GAACCAGGAAAGTAACCGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGCTAGAAAA 60

Qy   61 CTCTGTCGGTACCAACCCCAGAGCGTTGAGAGCAGCAGCCCAACCTCCACGCTTCCTTAACGGA 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTCTGTCGGTACCAACCCCAGAGCGTTGAGAGCAGCAGCCCAACCTCCACGCTTCCTTAACGGA 120

Qy  121 GAGGTGCAGGACTCAGACTTCACCAGCCACTCGGTCCCAGCCCTTGTACGCAAAGAGACG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GAGGTGCAGGACTCAGACTTCACCAGCCACTCGGTCCCAGCCCTTGTACGCAAAGAGACG 180

Qy  181 CCAAGGACGCGCGCTCTCCCGCGCGTCCAGGCAGCCCCAGCTTGCTGGCTTGCTCCGCCTG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCAAGGACGCGCGCTCTCCCGCGCGTCCAGGCAGCCCCAGCTTGCTGGCTTGCTCCGCCTG 240
```

RESULT 2
US-10-643-795A-5
; Sequence 5, Application US/10643795A
; Publication No. US20040241703A1
; GENERAL INFORMATION:
; APPLICANT: FREDERIC J. DESAUVAGE
; APPLICANT: GRETCHEN FRANTZ
; APPLICANT: KENNETH J. HILLAN
; APPLICANT: PAUL POLAKIS
; APPLICANT: ANDREW POLSON
; APPLICANT: VICTORIA SMITH
; APPLICANT: SUSAN D. SPENCER
; APPLICANT: THOMAS D. WU
; APPLICANT: ZEMIN ZHANG
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5026R1-US
; CURRENT APPLICATION NUMBER: US/10/643,795A
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 60/413,192
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US 60/419,008
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/426,847
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/484,959
; PRIOR FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 5
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 2
; OTHER INFORMATION: Unknown base
US-10-643-795A-5

```
Query Match             100.0%;  Score 1421;  DB 8;  Length 1450;
Best Local Similarity   100.0%;  Pred. No. 0;  Mismatches 0;
Matches 1421;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 GACCAGGAAAGTAAACGCTACAGACACGTGAGAGAATAGTTTCGCTCGGCGGCTAGAAAA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9 GACCAGGAAAGTAAACGCTACAGACACGTGAGAGAATAGTTTCGCTCGGCGGCTAGAAAA 68

Qy   61 CTCTGTCGGTACCAACCCCAGAGCGTTGAGAGCAGCCCACCTCCAGCGTTCCTTAACGGA 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   69 CTCTGTCGGTACCAACCCCAGAGCGTTGAGAGCAGCCCACCTCCAGCGTTCCTTAACGGA 128

Qy  121 GAGGTGCAGGACTTCACCAGCCCAGTCGGTGAGAGCAGCCCACCTCGTACGGCAAAGACG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  129 GAGGTGCAGGACTTCACCAGCCCAGTCGGTGAGAGCAGCCCACCTCGTACGGCAAAGACG 188

Qy  181 CCAAGGACGGCGCTCTCCCGCGTCCAGGCAGCAGCCCAGCCTTGCTTGGCTTGTGGCTG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  189 CCAAGGACGGCGCTCTCCCGCGTCCAGGCAGCAGCCCAGCCTTGCTTGGCTTGTGGCTG 248

Qy  241 CGTGCAGCACTCGGCCCGCGTGCAGCATGACCCTGTGGGAACGGCGTACTGCCTTTTTACC 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CGTGCAGCACTCGGCCCGCGTGCAGCATGACCCTGTGGGAACGGCGTACTGCCTTTTTACC 300
```

Qy 241 CGTGCAGCACTCGGCCCGCGTGCAGCATGACCCTGTGGGAACGGCGTACTGCCTTTTTACC 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 CGTGCAGCACTCGGCCCGCGTGCAGCATGACCCTGTGGGAACGGCGTACTGCCTTTTTACC 300

Qy 301 CCCAGCCCCCGGCCATGCCGCAGGCTTCCAGTGCCTTCAGCGCATCCCGGTTATTCTAGTGTTTT 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CCCAGCCCCCGGCCATGCCGCAGGCTTCCAGTGCCTTCAGCGCATCCCGGTTATTCTAGTGTTTT 360

Qy 361 TGGGCTTCTACGACAAGCTTCCTGCTCCTCTCATCTTGCCGGGGGATCCGGCCACTCGCGCTCGT 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 TGGGCTTCTAGGACAAGCTTCCTGCTCCTCTCATCTTGCCGGGGGATCCGGCCACTCGGCTCGGT 420

Qy 421 TTTGGTTGGTGAGAGTTCTTCTCAGTCTCGTTCAATAGGCCCGAGAAATTGTGGGCTTGTGCACT 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 TCACAGGGACCCCAGTGCATCAGCTGAACGAGGACCCATTGACTACTACAACGACGCAGTTCACCT 480

Qy 481 TCAGTGCAGAATGGTTCGTCGTTGGGTACAGTGAACACCAACACATCCTTACAAAGCCTTCAGGG 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 TCAGTGCAGAATGGTTCGTCGTTGGGTACAGTGAACACCAACACATCCTTACAAAGCCTTCAGGG 540

Qy 541 CACGGCGCCGTTACAGCCCCTGTGTCGGTCTCGTCTTCGTTCGGGGGCCTGGAGGGCATTAATATTACAC 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 CAGCGGCGGTTACAGCCCGTTGCTGGCGTCTCGTTCGGGCCTTGGAGGGCATTAATATTACAC 600

Qy 601 TCACAGGGGACCCCAGTGCATCAGCTGAACGAGGACCATTGACTACTACAACGAGCAGTTCACCT 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 GGCGTCTGAAAGAGAATTACGCGCCCGGAGTACGGCCAACGGCACTCGGAGAAATTGTGGCTGTGCACT 660

Qy 661 GGCGTCTGAAAGAGAATTACGCGCCCGGAGTACGGCCAACGGCACTCGGAGAAATTGTGGCTGTGCACT 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 GGGCTGAAAGAGAATTACGCGCCCGGAGTACGGCCAACGGCACTCGGAGAAGGCGCTGCCGG 720

Qy 721 ACCCAGTGCTCTACCTGGGCGGAGAAGTTCACACGCGAGTAGCCCTTGCGGCCCTGTACCACC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 ACCCAGTGCTCTACCTGGGCGGAGAAGTTCACACCGAGTAGCCCTTGCGGCCCTGTACCACC 780

Qy 781 AGTACCACCTGGGCGGGGACACTAGCCCTCGGCCATCGGGGTTTCTGCTTCTCGGC 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 AGTACCACCTGGGCGGGGACACTAGCCCTCGGCCATCGGGGTTTCTGCTTCTCGGC 840

Qy 841 TCCTTCTCCAACGTGCTCTGCTCCAGCGCCAGGCCTGGCACTGCTGA 900
        |||||||||||||||||||||||||||||||||||||||||||||||
Db 841 TCCTTCTCCAACGTGCTCTGCTCCAGCGCCAGGCCTGGCACTGCTGA 900

Qy 901 CCACCGGAGCCTTCGCGGGTCTTCGGCCTCTGGGCCTCCCATCTCTCTAGCGTGCCGC 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 CCACCGGAGCCTTCGCGGGTCTTCGGCCTCTGGGCCTCCCATCTCTCTAGCGTGCCGC 960

Qy 961 TCTGCCCGCGTCCGCCTAGGCTCTCTCCGCGCTCCACCACTCAGTACGGCGCCGCGCTTCTGGG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 TCTGCCCGCGTCCGCCTAGGCTCTCTCCGCGCTCCACCACTCAGTACGGCGCCGCGCTTCTGGG 1020

Qy 1021 TCACGCTGCGCAACCGGCGGCTTCGTCCTGTGGCCTTCTTCCTCCGGAGCGGCGGCCTAGAAAACTT 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 TCACGCTGGCAACCGGCGGCTTCGTCCTGTGGCCTTCTTCCTCCGGAGCGGCGGCCTAGAAAGT 1080

Qy 1081 ATGTTCGGCGCCCAGGCGCTCTTGCGACCCCTTCTGGACCAAAAGCGGCCAAGGACTGCAGCCAGG 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 ATGTTCGGCGCCCAGGCGCTCTTGCGACCCTTTCTGGACCAAAGCGGCCAAAGGCGACTGCAGCCAGG 1140

Qy 1141 AGAGAGGGGGGCTCACCTCCTTATCCTCGCGCCACTGCACAAGCACAGGCCGCTCTCCCAG 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 AGAGAGGGGGGCTCACCTCCTTATCCTCGCGCCACTGCACAAGCACAGGCCGCTCTCCCAG 1200

Qy 1201 ACTTAAAAATGTATCACCACTAACCTGTGAGGGGGACCCAATCTGGACTCCCTTCCCCCGCCT 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 ACTTAAAAATGTATCACCACTAACCTGTGAGGGGGACCCAATCTGGACTCCCTTCCCCCGCCT 1260

Qy 1261 TGGGACATCGCAGGCCCGGGAAGCAGTGCCGCGCCAGGCCTGGGCCGCCAGGAGAGCTCCAGGAA 1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 TGGGACATCGCAGGCCCGGGAAGCAGTGCCGCGCCAGGCCTGGGCCGCCAGGAGAGCTCCAGGAA 1320

Qy 1321 GGGCACTGAGCGCTGCTGGGCGTCGGGCGGAGGCCTGGGACATCCGGAGGGACCAGGGGAAAGTCTCC 1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 GGGCACTGAGCGCTGCTGGGCGTCGGGCGGAGGCCTGGGACATCCGGAGGGACCAGGGGAAAGTCTCC 1380

Qy 1381 TGGGGCGGATCTGTAAATAAACCTTTTTTTCTTTTTTGTTTTTT 1421
        ||||||||||||||||||||||||||||||||||||||||||||
Db 1381 TGGGGCGGATCTGTAAATAAACCTTTTTTTCTTTTTTGTTTTTT 1421

RESULT 3
US-10-948-518-5
; Sequence 5, Application US/10948518
; Publication No. US20050044492A1
; GENERAL INFORMATION:
; APPLICANT: FREDERIC J. DESAUVAGE
; APPLICANT: GRETCHEN FRANTZ
; APPLICANT: KENNETH J. HILLAN
; APPLICANT: PAUL POLAKIS
; APPLICANT: ANDREW POLSON
; APPLICANT: VICTORIA SMITH
; APPLICANT: SUSAN D. SPENCER
; APPLICANT: THOMAS D. WU
; APPLICANT: ZEMIN ZHANG
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5026R1-US
; CURRENT APPLICATION NUMBER: US/10/948,518
; CURRENT FILING DATE: 2004-09-22
; PRIOR APPLICATION NUMBER: US/10/643,795
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 60/413,192
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US 60/419,008
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/426,847
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/484,959
; PRIOR FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 5
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 2
; OTHER INFORMATION: Unknown base
US-10-948-518-5

Query Match   100.0%; Score 1421; DB 9; Length 1450;
Best Local Similarity   100.0%;
Matches 1421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 GAACCAGGAAAGTAACGGTACGGTACAGACAGCAGTGAGAAATAGTTTCGGCTCGGCCTAGAAAAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     9 GAACCAGGAAAGTAACGGTACGGTACAGACAGCAGTGAGAAATAGTTTCGGCTCGGCCTAGAAAAA 68

Qy    61 CTCTGTTCGGTTACCAACCCCAGAGCGTTGAGGAGCAGCCCACCTCCAGCCTTCCTTAACGGA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    69 CTCTGTTCGGTTACCAACCCCAGAGCGTTGAGGAGCAGCCCACCTCCAGCCTTCCTTAACGGA 128

Qy   121 GAGGTGCAGGACTCAGACTTCACCGACGTCCAGCCTTGTTACGGCAAAGAGACG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   129 GAGGTGCAGGACTCAGACTTCACCGACGTCCAGCCTTGTTACGGCAAAGAGACG 188

Qy   181 CCAAGGACGCGCCTCTCCCGCCTCCAGGCGCCAGCCCACTTGCTTGGCTTGGCCTG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   189 CCAAGGACGCGCCTCTCCCGCCTCCAGGCGCCAGCCCACTTGCTTGGCTTGGCCTG 248

Qy   241 CGTGCAGCACTTCGGCCGCGGTTGCAGCAGATGACCCTGTGGAACGGCGTACTGCCTTTTTACC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   249 CGTGCAGCACTTCGGCCGCGGTTGCAGCAGATGACCCTGTGGAACGGCGTACTGCCTTTTTACC 308
```

```
Db   249 CGTGCAGCACTTCGGCCGCGGCGCGTGGCTGTGAGCATGACCCTGTGGAACGGCGTACTGCCTTTTTACC 308

Qy   301 CCCAGCCCCGGGCCGCCAGGCCTTCAGGTTCCACTGCTCCTCATTCGTTATTCTTAGTGTTT 360
Db   309 CCCAGCCCCGGGCCCGCCATGCCCGGCCTTCAGGTTCCACTGCTCCTCATCGTTATTCTTAGTGTTT 368

Qy   361 TGGCTCTTAGCAGCAGCAAGCTTCCTCCTCTCATTTGCGGGGGATCCGTGGCCACTCCGCGCTGT 420
Db   369 TGGCTCTTAGCAGCAGCAAGCTTCCTCCTCTCATTTGCGGGGGATCCGTGGCCACTCCGGCTGGT 428

Qy   421 TTTGGTTGGTGAGAGTTCTTCTCAGTCTCGTTCATAGGCCCGAGAAATTGTGGGCTTGTGCACT 480
Db   429 TTTGGTTGGTGAGAGTTCTTCTCAGTCTCGTTCATAGGCCCGAGAAATTGTGGGCTTGTGCACT 488

Qy   481 TCAGTGCAGAATGGTTCGTTCGGGTACAGTGAACACCAACACATCCTTACAAAGCCCTTCAGCG 540
Db   489 TCAGTGCAGAATGGTTCGTTCGGGTACAGTGAACACCAACACATCCTTACAAAGCCCTTCAGCG 548

Qy   541 CAGCGCCGGTTACAGCCCGTTGGTTCGTGTCGGTTCCGTGGGCCTGGAGGGGCATTAATATTCAC 600
Db   549 CAGCGGCGCTTACAGCCCGTTCGGTTCTCGTGTCGGTTCCGTGGGCCTGGAGGGGCATTAATATTACAC 608

Qy   601 TCACAGGGACCCCAGTCGCATCAGTGAACGAGAGACCATTGACTACTACAAACGAGCAGTTCACTT 660
Db   609 TCACAGGGACCCCAGTGCATCAGTGACGAGAGACCATTGACTACTACAAACGAGCAGTTCACT 668

Qy   661 GGCCTCTGAAAGAGAATTACGCCCGCGGAGTACGGCGACACGGCAGCGAAGAGGGGCTGCCGG 720
Db   669 GGCCTCTGAAAGAGAATTACGCCCGCGGAGTACGGCGACACGGCAGCGAAGAGGGGCTGCCGG 728

Qy   721 ACCCAGTGCTCTACCTGGGGGAGAAGTTCACACGCGGAGTAGCCCTGTACCACC 780
Db   729 ACCCAGTGCTCTACCTGGGGGAGAAGTTCACACGCGGAGTAGCCCTGTACCACC 788

Qy   781 AGTACCACCTGGGCGGGACTACAGCCCTCGGCCTATGGGGTTCCGTTCTTCTGGC 840
Db   789 AGTACCACCTGGGCGGGACTACAGCCCTCGGCCTATGGGGTTCCGTTCTTCTGGC 848

Qy   841 TCCTCTCAAGCTGCTCTCCAGCGCCGCCCCCCGCCTCTACGGAGCCACTGCTGA 900
Db   849 TCCTCTCAAGCTGCTCTCCAGCGCCGCCCCCCGCCTCTACGGAGCCACTGCTGA 908

Qy   901 CCACCGGAGCCCTTCGGCGCTCTTCGGGTGTTCTTCGCGCTGTTGGCCGCCGC 960
Db   909 CCACCGGAGCCCTTCGGCGCTCTTCGGGTGTTCTTCGCGCTGTTGGCCGCCGC 968

Qy   961 TCTGCCCGGTCCGGCCTAGGCCTCCTCCGCGCGCCTCACCACTCAGTACGTACGGGG 1020
Db   969 TCTGCCCGGTCCGGCCTAGGCCTCCTCCGCGCGCCTCACCACTCAGTACGTACGGGG 1028

Qy  1021 TCACGCTGCAACCGGCGTCCCTGTTCTTCGGGTGTGGTGAGTCTCCAGT 1080
Db  1029 TCACGCTGCAACCGGCGTCCCTGTTCTTCGGGTGTGGTGAGTCTCCAGT 1088

Qy  1081 ATGTTCGGCCCCAGCGCTCTTGCACACCCTTCTGACGAACAAGGCGCTGCTGCAGGA 1140
Db  1089 ATGTTCGGCCCCAGCGCTCTTGCACACCCTTCTGACGAACAAGGCGCTGCTGCAGGA 1148

Qy  1141 AGAGAGGGGGGGGCCTCACCTCTTATCCTCGGCGGGACCCACTGCACAAGCACCGGCCG 1200
Db  1149 AGAGAGGGGGGGCTCGACCTCTTATCCTCGGACCCACTGCACAAGCACCGGCCG 1208

Qy  1201 ACTTAAAAATGTATCACCTAACCTGTGAGGGGACCCAAATCTGGACTCTCCCGCCT 1260
Db  1209 ACTTAAAAATGTATCACCACTAACCTGTGAGGGGACCCAAATCTGGACTCTTCCCGCCT 1268

Qy  1261 TGGGACCATCCGCAGGCGCGGGGAAGCAGTGCCCCGCCTGGGCCTTCAGGGAA 1320
Db  1269 TGGGACCATCCGCAGGCGCGGGGAAGCAGTGCCCCGCCTGGGCCTTCAGGGAA 1328

Qy  1321 GGGGCACTGAGGGCTGCTGGGCGGAGGCCTCGGACATCCGCGAGCGCCAGGGGAAAGTCTCC 1380
Db  1329 GGGGCACTGAGGGCTGCTGGGCGGAGGCCTCGGACATCCGCGAGCGCCAGGGGAAAGTCTCC 1388

Qy  1381 TGGGGCGATCTGTAAATAAACCTTTTTTTGTTTTT 1421
Db  1389 TGGGGCGATCTGTAAATAAACCTTTTTTTGTTTTT 1429
```

RESULT 4
US-10-264-237-1097
Sequence 1097, Application US/10264237
Publication No. US20040009491A1

; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: Patentin Ver. 3.1
;
; SEQ ID NO 1097
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-237-1097

Query Match   99.8%;   Score 1418.6;   DB 6;   Length 1474;
Best Local Similarity   99.8%;
Matches 1418;   Conservative   2;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 GAACCAGGAGAAAGTAACGGCTACAGACACAGTGAGAAATAGTTCGGTTCGCGGGGCTAGAAAAA    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     19 GAACCAGGAGAAAGTAACGGCTACAGACACAGTGAGAAATAGTTCGGTTCGCGGGGCTAGAAAAA    78

Qy     61 CTCTGTCGGTACCAACCCCGAGAGGGTTGAGAGCAGCCCACCTCCACGCTTCCTTAACGGA   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 CTCTGTCGGTACCAACCCCGAGAGGGTTGAGAGCAGCCCACCTCCACGCTTCCTTAACGGA   138

Qy    121 GAGGTGCAGGACTCAGACTTCACCAGGCCCACTCCAGCCCTTGTACGGCAAAGAGACG   180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 GAGGTGCAGGACTCAGACTTCACCAGGCCCACTCCAGCCCTTGTACGGCAAAGAGACG   198

Qy    181 TCAAGGACGGCGCTCTCCCGCGTCCAGGCAGCCCCAGCGTTGGCTTGGCCTTGCCTGG   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 TCAAGGACGGCGCTCTCCCGCGTCCAGGCAGCCCCAGCGTTGGCTTGGCCTTGCCTGG   258

Qy    241 CGTTGCAGCACTCGGCCGCGGTGCGAGCAGTGAACGGGGTACGCCTTTTTACC   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 CGTTGCAGCACTCGGCCGCGGTGCGAGCAGTGAACGGGGTACGCCTTTTTACC   318

Qy    301 CCCAGCCCGGGCCATGCCGCGGCAGGGCTTCAGGGTTCCACTGCTGGT   360
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    319 CCCAGCCCGGGCCATGCCGCGGCAGGGCTTCAGGGTTCCACTGCTGGT   378

Qy    361 TGGCTCTAGCAGCAGAAGCTTCCTGCTCATCTTGCGGGAGATCCGGTGGCACT   420
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    379 TGGCTCTAGCAGCAGAAGCTTCCTGCTCATCTTGCGGGAGATCCGGTGGCACT   438

Qy    421 TTTGGTTGGTGAGAGTTCTTCTCAGTGTTCATAGCGGCCAGAAAATTGTGGGCTGCACT   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    439 TTTGGTTGGTGAGAGTTCTTCTCAGTGTTCATAGCGGCCAGAAAATTGTGGGCTGCACT   498

Qy    481 TCAGTGCAGAATGGTTCGTGGGTACAGTGAACACCAACACATCCTACAAAAGCCTTCAGCG   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    499 TCAGTGCAGAATGGTTCGTGGGTACAGTGAACACCAACACATCCTACAAAAGCCTTCAGCG   558

Qy    541 CAGGGCGCGTTACAGCCCGTGTCGGTCTGCGGCCTTACAGGGGCATTAATTATTACAC   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    559 CAGGGCGCGTTACAGCCCGTGTCGGTCTGCGGCCTTACAGGGGCATTAATTATTACAC   618

Qy    601 TCACAGGGACCCCAGTGCATCAGTTGAACGAGGAGACCCATTGGACTACAAGGAGCAGTTCACCT   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCACAGGGACCCCAGTGCATCAGTTGAACGAGGAGACCCATTGGACTACAAGGAGCAGTTCACCT   660
```

(Alignment continues — query/database sequence pairs from positions 301 through 1421/1429)

```
Qy   1381 TGGGGGCGATCTGTAAATAAAACCTTTTTTTTTTTTTTGTTTTT   1421
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1389 TGGGGGCGATCTGTAAATAAAACCTTTTTTTTTTTTTTGTTTTT   1429
```

Db 619 TCACAGGGACCCCAGTGCATCAGTGACGAGACCATTGACTACTACAACGAGCAGTTCACTT 678
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 661 GGCGTTCTGAAAGAGAATTAGCCCGGGGAGTAGCCGAACGCACTGGAGAAGGGGCTGCCGG 720

Db 679 GGGCGTCTGAAAGAGAATTAGCCCGGGGAGTAGCCGAACGCACTGGAGAAGGGGCTGCCGG 738
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 721 ACCCAGTGCTCTACCTGGGCGGGAGAGTTCACACCGAGTAGCCCTTGCGCGCCTGTACCACC 780

Db 739 ACCCAGTGCTCTACCTGGGCGGGAGAGTTCACACCGAGTAGCCCTTGCGCGCCTGTACCACC 798
         ...

Qy 1381 TGGGGCGATCTGTAAATAAACCTTTTTTTCTTTTGTTTTTT 1421
         |||||||||||||||||||||||||||||||||||||||||
Db 1399 TGGGGCGATCTGTAAATAAACCTTTTTTTTCTTTTGTTTTTT 1439

RESULT 5

SEQ ID NO 4
LENGTH: 1420
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20030068311A1 3221661CA2
US-10-187-657-4

Query Match   99.6%; Score 1415.8; DB 5; Length 1420;
Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;
Matches 1417; Conservative 0; Mismatches 0;

Qy   1 GAACCAGGAAAGTAACGGCTACAGACAGCAGTGAGAAATAGTTTCGCTCGGCCGGCTAGAAAAA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GAACCAGGAAAGTAACGGCTACAGACAGCAGTGAGAAATAGTTTCGCTCGGCCGGCTAGAAAAA 60

...

Qy 841 TCCTTCTCCAAGTGCTGCTCTCCACCGGCGGCCCCCGGCTCTACGGAGGCCTGGCACTGCTGA 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 TCCTTCTCCAAGTGCTGCTCTCCACCGGCGGCCCCCGGCTCTACGGAGGCCTGGCACTGCTGA 900

RESULT 6
US-10-773-236-187
; Sequence 187, Application US/10773236
; Publication No. US20050208602A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et. al.
; TITLE OF INVENTION: 89 Human Secreted Proteins
; FILE REFERENCE: PS751P1
; CURRENT APPLICATION NUMBER: US/10/773,236
; CURRENT FILING DATE: 2004-02-09
; PRIOR APPLICATION NUMBER: 60/311,085
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,209
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/US02/25107
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/330,629
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: PCT/US02/33985
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 60/331,046
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: PCT/US02/35606
; PRIOR FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 60/358,554
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: PCT/US03/04819
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/358,714
; PRIOR FILING DATE: 2002-02-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 396
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 187
; LENGTH: 1491
; TYPE: DNA

; ORGANISM: Homo sapiens
US-10-773-236-187

Query Match 99.4%; Score 1412; DB 9; Length 1491;
Best Local Similarity 99.9%; Pred No. 0;
Matches 1412; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy   8 GAAAGTAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGCTAGAAAAACTCTGTC 67
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43 GAAAGTAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGCTAGAAAAACTCTGTC 102

Qy   68 GGTACCAACCCAGAGACGGTTGAGAGCAGCAGCCCACCTCCAGCCTTCCTTAACGGAGGA 127
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   103 GGTACCAACCCAGAGACGGTTGAGAGCAGCAGCCCACCTCCAGCCTTCCTTAACGGAGGA 162

Qy   128 AGGACTCAGACTTCACCAGCCCCACTCCGGTCCCCAGCCTTGTACGCAAAGAAGACGCAAGGA 187
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   163 AGGACTCAGACTTCACCAGCCCCACTCCGGTCCCCAGCCTTGTACGCAAAGAAGACGCAAGGA 222

Qy   188 CGCGCTCTCCCGGCTCCAGGCAGCCCCAGCTTGCTTGGCTTGCCCCGCCTCGGTGTGCAG 247
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   223 CGCGCTCTCCCGGCTCCAGGCAGCCCCAGCTTGCTTGGCTTGCCCCGCCTCGGTGTGCAG 282

Qy   248 CACTCGGCCCGCGGTGCAGCATGACCGTGTGGAACGGGTACTGCCTTTTTACCCCCAGCC 307
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   283 CACTCGGCCCGCGGTGCAGCATGACCGTGTGGAACGGGTACTGCCTTTTTACCCCCAGCC 342

Qy   308 CCGGCCATGCCGCAGGCTTCACGTCCACTGCTCATCGTTATTCTTAGTGTTTGTTTTTCT 367
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   343 CCGGCCATGCCGCAGGCTTCACGTCCACTGCTCATCGTTATTCTTAGTGTTTTTGGCTCT 402

Qy   368 AGCAGCAAGCTTCCTTGCTCATCTTGCCCGGGGGATCCGGTAGGCCCACTCGCGCTGGTTTTGGTT 427
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   403 AGCAGCAAGCTTCCTTGCTCATCTTGCCCGGGGGATCCGGTAGGCCCACTCGCGCTGGTTTTGGTT 462

Qy   428 GGTGAGAGGTTCTTCAGTCTGTTCGTTCATAGGCGCAGAAATGTGGCGTGGCACTTCAGTGC 487
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   463 GGTGAGAGGTTCTTCAGTCTGTTCGTTCATAGGCGCAGAAATGTGGCGTGGCACTTCAGTGC 522

Qy   488 AGAATGGTTCGTGGGTACAGTGAACACAAACCTTCAGCGCAGCGCGCG 547
               ||||||||||||||||||||||||||||||||||||||||||||||||
Db   523 AGAATGGTTCGTGGGTACAGTGAACACAAACCTTCAGCGCAGCGCGCG 582

Qy   548 CGTTACAGCCCGTGTCCGGTCTGCTCCGTGGCGGCCTGGAGGCGCATTAATATTACACTCACAGG 607
               ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db   583 CGTTACAGCCCGTGTCCGGTCTGCTCCGTAGGCGCATTAATATTACACTCACAGG 642

Qy   608 GACCCCAGTGCATCAGCTGAACGAGACCATTGACTACAACGACAGTTCACCTGGCGTCT 667
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   643 GACCCCAGTGCATCAGCTGAACGAGACCATTGACTACAACGACAGTTCACCTGGCGTCT 702

Qy   668 GAAAGAGAATTACCCGCGGAGTACGCGAACGCACTGAGAGAGGGGGCTGCCGGGACCCAGT 727
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   703 GAAAGAGAATTACCCGCGGAGTACGCGAACGCACTGAGAGAGGGGGCTGCCGGGACCCAGT 762

Qy   728 GCTCTACCTGGGCGGAGAAGTTCACACCCGACGTAGGCCCTTGGGGCTTGTACCACCAGTACCA 787
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   763 GCTCTACCTGGGCGGAGAAGTTCACACCCGACGTAGGCCCTTGGGGCTTGTACCACCAGTACCA 822

Qy   788 CCTGGCGGGACACTACGCCTCGGCCACGCTATGGGGTGGCGTTCTGCTTCTGGCTCCTCTC 847
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   823 CCTGGCGGGACACTACGCCTCGGCCACGCTATGGGGTGGCGTTCTGCTTCTGGCTCCTCTC 882

Qy   848 CAACGTGCTGCTCTCCACCCCGCGGCTCTACAGGAGCGGCCTGGCACTGCACTCTGACCACCCGG 907
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   883 CAACGTGCTGCTCTCCACCCCGCGGCTCTACAGGAGCGGCCTGGCACTGCACTCTGACCACCCGG 942

Qy   908 AGCCTTCGCGCTCTTCGCGGGCGTCTTCGGCCTTCCATCTCTACGCGGTCGCCGGCTCTGCC 967
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   943 AGCCTTCGCGCTCTTCGCGGGCGTCTTCGGCCTTCCATCTCTACGCGGTCGCCGGCTCTGCC 1002

Qy   968 GCTCCGCCTAGGCTCCCCGCGCTCACCACTACGGCGGCCGCCGCTTCTGGGTCACGCT 1027
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GCTCCGCCTAGGCTCCCCGCGCTCACCACTACGGCGGCCGCCGCTTCTGGGTCACGCT ...

Qy   961 TCTGCCCGGCTCCCTCCCTAGGCTCCTCCGCGGGCTCACCCACTTCAGTAGCTTCTGG 1020
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TCTGCCCGGCTCCCTCCCTAGGCTCCTCCGCGGGCTCACCCACTTCAGTAGCTTCTGG 1020

Qy   1021 TCACGCTGGCAACCGGCCGTCCTGTGCCTCTCCTCGGAGGGGGGCGGGCGTGGTGGAGTCTCCAGT 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCACGCTGGCAACCGGCCGTCCTGTGCCTCTCCTCGGAGGGGGGCGGGCGTGGTGGAGTCTCCAGT 1080

Qy   1081 ATGTTCGGCCCAGCGCTTCTTCGGACACCTTCTCCTTCCTCGGAGGACTGCAGCCCAGT 1140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ATGTTCGGCCCAGCGCTTCTTCGGACACCTTCTCCTTCCTCGGAGGACTGCAGCCCAGT 1140

Qy   1141 AGAGAGGGGGGGCTCACCTCTTATCCTCGGCGACCCACTCCACAAGCCACAAAGGGCCAAGGA 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AGAGAGGGGGGGCTCACCTCTTATCCTCGGCGACCCACTCCACAAGCCACAAAGGGCCAAGGA 1200

Qy   1201 ACTTAAAATGTATCACCACTAACCTGTGAGGGGGGAGACCCAAATCGACTCCTTCCCTCCCGGCCT 1260
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACTTAAAATGTATCACCACTAACCTGTGAGGGGGGAGACCCAAATCGGACTCCTTCCCGCGCCT 1260

Qy   1261 TGGGACATCGCACGCCGGGAAGCAGTGCCCGCGCAGGCCTGGGAACGGGTACTGCCTTTTTACCCCCAGAA 1320
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGGGACATCGCACGCCGGGAAGCAGTGCCCGCGCAGGCCTGGGAACGGGTACTGCCTTTTTACCCCCAGAA 1320

Qy   1321 GGGCACTGAGCGCTGCTGGCGCGAGACATCCGCAGCACCAGGGAGAGAAAAGTCTCC 1380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GGGCACTGAGCGCTGCTGGCGCGAGACATCCGCAGCACCAGGGAGAGAAAAGTCTCC 1380

Qy   1381 TGGGGCGGATCTGTAAAATAAAACCTTTTTTTCTTTTGTTTT 1419
               ||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGGGCGGATCTGTAAAATAAAACCTTTTTTTCTTTTGTTTT 1419

Qy   901 CCACCGGAGCCTTCGCGCTCTTCGGGGGTCTTCGCCTTCCCATCTTCTAGGGTGCCGC 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CCACCGGAGCCTTCGCGCTCTTCGCGGGGTCTTCGCCTTCCCATCTTCTAGGGTGCCGC 960

RESULT 7
US-10-187-657-9
Sequence 9, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 1594
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 110769_Mm.1
US-10-187-657-9

Query Match   48.8%;   Score 693; DB 5; Length 1594;
Best Local Similarity 75.9%;   Pred. No. 4.2e-204;
Matches 914;   Conservative   0; Mismatches 275; Indels 16; Gaps 4;

Qy   227 CCTGGTGTTCTTGTTGTTGGCTGGAATCGGTGGCCA 286
          ||||| |||||||||||||| |||||| |||||||
Db   1304 TTTTT 1308

[Sequence alignment block — Qy (query) and Db (database) rows with match bars, columns spanning positions 1003–1594 / 107–1421 as listed in the original]

RESULT 8
US-10-187-657-5
; Sequence 5, Application US/10187657
; Publication No. US20030683111A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
;
; SEQ ID NO 5
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030683111A1 32216611R6
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 455, 480, 483
; OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-5

Query Match 32.3%; Score 459.6; DB 5; Length 522;
Best Local Similarity 96.8%; Pred. No. 8.6e-132;
Matches 510; Conservative 0; Mismatches 12; Indels 5; Gaps 4;

```
Qy      1 GAACCAGGAAAGTAAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGGCCGGCTAGAAAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAACCAGGAAAGTAAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGGCCGGCTAGAAAA 60

Qy     61 CTCTGTCCGGTACCAAACCCCAGAGACGGTTGAGAGCAGCCCACTCCAGCGCTTCCTTAAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     61 CTCTGTCCGGTACCAAACCCCAGAGACGGTTGAGAGCAGCCCACTCCAGCGCTTCCTTAACGGA 120

Qy    121 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCGTTGTACGCAAAGAGACG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCGTTGTACGCAAAGAGACG 180

Qy    181 CCAAGGACGCGCTTCTCCCGCGTCCAGCCCCAGCGTTGCTGGCTTGCCTGCCGCCCTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCAAGGACGCGCTTCTCCCGCGTCCAGCCCCAGCGTTGCTGGCTTGCCTGCCGCCCTG 240

Qy    241 CGTGCAGCACTCCGGCCGGCGTGCAGCATGACCCGTGGAACGCGTACTACTGCCCTTACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGTGCAGCACTCCGGCCGGCGTGCAGCATGACCCGTGGAACGCGTACTACTGCCCTTACC 300

Qy    301 CCCAGCCCCGGCATGCCCGGCGCTTCAGCGTTCCACTCGTCAGCGTTATTCTAGTGTTTT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCCAGCCCCGGCATGCCCGGCGCTTCAGCGTTCCACTCGTCAGCGTTATTCTAGTGTTTT 360

Qy    361 TGGCTCTAGCAGCAAGCTTCCTTGCTCCTGCGGGGATCCGTGGGCCACTCGCGCTGGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGG-TCTAGCAGCAAG--TTCCTGCTCCTGCGGGGATCCGTGGGCCACTCGCGCTGGT 417

Qy    421 TTTGGTTGGTGAGAGTTCTTCTCAGTCTGTTCATGAGGNGCAGAAAATTGTGGGCTGTGCACT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    418 TTTGGTTGGTGAGAGTTCTTCTCAGTCTGTTCATGAGGNGCAGAAAATTGTGGGCTGTGCACT 477

Qy    481 TCAGTGCAGAATGGTTCGTGGGGTACAGTGAAACACCAACACATCCTAC 527
          ||||||||||||||||||||||||||||||||||||||||||||
Db    478 TCNGT-NAGAATGGTTCGTGGGGTACATTG-ACACCAACACATCTAC 522
```

RESULT 9
US-10-773-236-69
; Sequence 69, Application US/10773236
; Publication No. US20050208602A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et. al.
; TITLE OF INVENTION: 89 Human Secreted Proteins
; FILE REFERENCE: PS751P1
; CURRENT APPLICATION NUMBER: US/10/773,236
; CURRENT FILING DATE: 2004-02-09
; PRIOR APPLICATION NUMBER: 60/311,085
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,209
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/US02/25107
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/330,629
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: PCT/US02/33985
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 60/331,046
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: PCT/US02/35606
; PRIOR FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 60/358,554
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: PCT/US03/04819
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/358,714
; PRIOR FILING DATE: 2002-02-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 396
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 69
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-773-236-69

Query Match 29.1%; Score 413; DB 9; Length 580;
Best Local Similarity 100.0%; Pred. No. 2.5e-117;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GAACCAGGAAAGTAAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGGCCGGCTAGAAAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAACCAGGAAAGTAAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGGCCGGCTAGAAAA 60

Qy      5 GAACCAGGAAAGTAAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGGCCGGCTAGAAAA 64
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5 GAACCAGGAAAGTAAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGGCCGGCTAGAAAA 64

Qy     61 CTCTGTCGGTACCAAACCCCAGAGACGGTTGAGAGCAGCCCACTCCAGCGCTTCCTTAACGGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65 CTCTGTCGGTACCAAACCCCAGAGACGGTTGAGAGCAGCCCACTCCAGCGCTTCCTTAACGGA 124

Qy    121 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCGTTGTACGCAAAGAGACG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    125 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGTCCCAGCGTTGTACGCAAAGAGACG 184

Qy    181 CCAAGGACGCGCTCTCCCGCGTTCCAGCCCCAGCGTTGCTGGCTTGCCTGCCGCCCTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    185 CCAAGGACGCGCTCTCCCGCGTTCCAGCCCCAGCGTTGCTGGCTTGCCTGCCGCCCTG 244

Qy    241 CGTGCAGCACTCGGCCGCGCTCTCCCGCGTGCAGCCCACGCCGCTTGCTGGCTTCAGTGTTACGCAAAGAGACG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    245 CGTGCAGCACTCGGCCGCGCTCTCCCGCGTGCAGCCCACGCCGCTTGCTGGCTTCAGTGTTACGCAAAGAGACG 304

Qy    301 CCCAGCCCCGGCCATGCCCGGCGCGAGGCTTCAGCGTTCCACTCGTTATTCTAGTGTTTTTACC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    305 CCCAGCCCCGGCCATGCCCGGCGCGAGGCTTCAGCGTTCCACTCGTTATTCTAGTGTTTTTACC 364

Qy    361 TGGCTCTAGCAGCAAGCTTCCTGCTCCTGCGGGGATCCGTGGGCCACTCG 413
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    365 TGGCTCTAGCAGCAAGCTTCCTGCTCCTGCGGGGATCCGTGGGCCACTCG 417
```

RESULT 10

US-10-450-763-2367/c
; Sequence 2367, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 2367
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (229)..(92)
; OTHER INFORMATION: 41% homologous to Homo sapiens Human secreted protein, SEQ ID
; OTHER INFORMATION: NO: 7589,accession number G03508,Smith-Waterman Score=71.
US-10-450-763-2367

Query Match   28.3%; Score 402.8; DB 9; Length 406;
Best Local Similarity 99.5%; Pred. No. 3.4e-114;
Matches 404; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   8 GAAAGTAACGGGCTACAGAGACAGAGTGAGAAATAGTTCGCTCGCAGAAAAACTCTGTC 67
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   406 GAAAGTAACGGCTACAGACAGTGAGAAATAGTTCGCTCGCAGAGAAAAACTCTGTC 347

Qy   68 GGTACCAACCCCAGAGAGCGTTGAGAGCAGCCCACCTCCACGCTTCCTTAACGGAGAGGTGC 127
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346 GGGACCAACCCCAGAGCGTTGAGAGCAGCCCACCTCCACGCTTCCTTAACGGAGAGGTGC 287

Qy   128 AGGACTCAGACTTCACCAGCCCACTCGGGTCCCAGCCTTGTACGCAAAGAGACGCCAAGGA 187
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286 AGGACTCAGACTTCACCAGCCCACTCGGGTCCCAGCCTTGTACGCAAAGAGAGACGTCAAGGA 227

Qy   188 CGGCGCTCTCCCGGCGTCCAGGCAGCCCACGTACTGCCTCCCCGCCTTGCTGCTCGCAG 247
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   226 CGGCGCTCTCCCGGCGTCCAGGCAGCCCACGTACTGCCTCGCCCCGCCTTGCTGCTCGCAG 167

Qy   248 CACTCGGCGCCGCGGTGCGAGCAATGACCCGTGGGAACGGGCGTACTGCCTCTTTTACCCCCAGCC 307
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   166 CACTCGGCGCCGCGGTGCAGCATGACCCGTGTGGAACGGGCGTACTGCCTTTTTTACCCCCAGCC 107

Qy   308 CCGGCCATGCGCGCAGCCTTCAGCGTTCCACGTTCCATCGTTATTCTAGTGTGTTTTTGGCTCT 367
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106 CCGGCATGCGCAGCCTTCAGCGTTCCACGTTCCATCGTTATTCTAGTGTTTTTGGCTCT 47

Qy   368 AGCAGCAAGCTTCCTGCTCATCTTGCCGGGGATCCGTGGCCACTCG 413
            ||||||||||||||||||||||||||||||||||||||||||||||
Db   46 AGCAGCAAGCTTCCTGCTCATCTTGCCGGGGATCCGTGGCCACTCG 1

RESULT 11
US-09-925-065A-72673/c
; Sequence 72673, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 10882.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096

; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72673
; LENGTH: 1376
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-72673

Query Match   27.9%; Score 396.6; DB 4; Length 1376;
Best Local Similarity 99.7%; Pred. No. 3.9e-112;
Matches 396; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   1 GAACCAGGAAAGTAACGGGCTACAGACAGTGAGAAATAGTTCGCTCGCAGCGGCTAGAAAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   397 GAACCAGGAAAGTAACGGGCTACAGACAGTGAGAAATAGTTCGCTCGCAGCGGCTAGAAAA 338

Qy   61 CTCTGTCGGTACCAACCCCAGAGAGCGTTGAGAGCAGCCCACCTCCACGCTTCCTTAACGGA 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   337 CTCTGTCGGTACCAACCCCAGAGAGCGTTGAGAGCAGCCCACCTCCACGCTTCCTTAACGGA 278

Qy   121 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGGTCCCAGCCTTGTACGCAAAGAGACG 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   277 GAGGTGCAGGACTCAGACTTCACCAGCCCACTCGGGTCCCAGCCTTGTACGCAAAGAGACG 218

Qy   181 CCAAGGACGGCGCTCTCCCGGCGTCCAGGCAGCCCACGTACTGCCTTGCTGCTCGCAGCCGCTG 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   217 YCAAGGACGGCGCCTCTCCCGGCGTCCAGGCAGCCCACGTTGCTGCTCGCAGCCGCTG 158

Qy   241 CGTGCAGCACTCGGCGCCGCGGTGCAGCATGACCCGTGGGAACGGGCGTACTGCCTTTTTACC 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157 CGTGCAGCACTCGGCGCCGCGGTGCAGCATGACCCGTGGGAACGGGCGTACTGCCTTTTTACC 98

Qy   301 CCCAGCCCCGGCCATGCGCAGCCTTCAGCGTTCCACTGCTCATCGTTATTCTAGTGTGTTT 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97 CCCAGCCCCGGCCATGCGCAGCCTTCAGCGTTCCACTGCTCATCGTTATTCTAGTGTGTTT 38

Qy   361 TGGCTCTAGCAGCAAGCTTCCTGCTCATCTTGCCGGGG 397
            ||||||||||||||||||||||||||||||||||||||
Db   37 TGGCTCTAGCAGCAAGCTTCCTGCTCATCTTGCCGGGG 1

RESULT 12
US-10-094-749-1195/c
; Sequence 1195, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, YOSHIKAZU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 08435/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1195
; LENGTH: 2684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-1195

Query Match   27.5%; Score 391; DB 6; Length 2684;
Best Local Similarity 97.5%; Pred. No. 2.5e-110;
Matches 397; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1015 TCTGGGTCACGCTGGCAACCGGGCGTCCGTGTCGTGGTGAGTC 1074
Db 2363 TTTCGATCCCACCGCCCACAGGCGGTCCTGTGCCGGTGGTGAGTC 2304

Qy 1075 TCCAGTATGTTCGGCCCAGCGCTCTTCCGCACCCTTCTCGGAGCACGACTGCA 1134
Db 2303 TCCAGTATGTTCGGCCCACGCGCTCTTCGCACCCTTCTCGGACCAAAGCGGCCAAGGACTCGCA 2244

Qy 1135 GCCAGGAGGAGAGAGGGGGCTCACTCTTCTTATCCTCGGCGACACCAAAGCACAGCAGGCCCGCTC 1194
Db 2243 GCCAGGAGAGAGAGGGGGCTCACCTTCTTATCCTCGGCGACACCCATGCACAAGCACAGCGCCCGCTC 2184

Qy 1195 TCCCAGGACTTAAAATGTATCACCAACTAACCTGTGAGGGGGGACCCAATCTGGACTCCTTCC 1254
Db 2183 TCCCAGACTTAAAATGTATCACCACTAACTGTGTGAGGGGGGACCCAATCTGGACTCCTTCC 2124

Qy 1255 CCCCCTTGGGACACATCGCAGGCCCCCGCGCCCAGGCGCTGGGCGTGGCCAGGAGAGCTC 1314
Db 2123 CCGCCTTGGGACACATCGCAGGCCCCGGGCCGGAAGCAGTGCCGCCAGGGGGTCAAGGTGCTGCTCTCCC 2064

Qy 1315 CAGGAAGGGCACTGAGCGCCTGCTCGTCGGCGCAGGGCCTCGGACATCGGCGCACCAGGGAA 1374
Db 2063 CAGGAAGGGCACTGAGCGCGTGCTCGGACGGCCTCGGACATCGGCGCACCAGGGAA 2004

Qy 1375 GTCTCCTGGGGCCGATCTGTAAATAAACCTTTTTTTTTTTTTTTTTT 1421
Db 2003 GTCTCCTGGGGCCGATCTGTAAATAAACCTTTTTTTTTTCTTTTTTGTTTTTT 1957

; RESULT 13
US-10-187-657-7
; Sequence 7, Application US/10187657
; Publication No. US20030683111A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No. US20030068311A1 SBQA0451OD1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 140, 165, 203, 237, 269, 289, 307, 338, 344, 347, 354, 390, 430, 445
; OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-7

Query Match   27.1%; Score 385.8; DB 5; Length 450;
Best Local Similarity 93.7%; Pred. No. 6.6e-109;
Matches 404; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

Qy 624 CTGAACGAGACCATTGACTACAACGACGAGCAGCAGTTCACCTGGCGTCTGAAAGAGGAATTACGCC 683
Db   21 CTGAACGAGACGCCACTGGACAACGGAGCAGTTCACCTGGCGTCTGAAAGAGGAATTACGCC 80

Qy 684 GCCGGAGTACGCGCAACGGCACTGGGAGGAAGAGGGGGGCTGCCGGGACCCAGTTGCTCTCTACCTGGGCGGAG 743
Db   81 GCCGGAGTACGCCGGAACGGCACTGGGAGGAAGAGGGGGGCTGCCGGGACCCAGTTGCTCTGAAAGAGGAATTACGCC 140

Qy 744 AAGTTCACACGACGTAGCCCTGGGCCTTGGCGCCCTGGGCCTTGCTGTGTCGGTCTTTGGTCTCTGCCAGTGGCGGGGGACACTAC 803
Db 141 AAGTTCACACGACGTAGCCCTTGCNGCCCTTGTGTCACCCACCAGTACCACCTGGCGTGCGGGGACACTAC 200

Qy 804 GCCTCGGCCACAGCGCTATGGGTGGGCGGTTCTGCTTCTGTCGTTCTGGCTCCTCTCCAAGTGCTGGCTGCTCTCC 863
Db 201 GCNTCGGCCACGGCTATGGGTGGGCGGTTCTGCTTCTGTCGTTCTGGCTCCTCTCCAAGTGCTGGCTGCTCTCC 260

Qy 864 ACGCGCGGCCCCGCTCTACGGAGGGCGTGGGCACTGCTGACCAGGCCGAGCCGTTCGGCGCTCTTC 923
Db 261 ACGCCGGCGCNCCGCTCTACGGAGGGCCTGGGNACTGCTGACCACCGGAGNCTTCGGCGCTCTTC 320

Qy 924 GGGGGTCTTCGGCCTTCCATCTCTAGGGTGCCGGCTCGGCCTTCGCCGCCCTAGGGCTCC 983
Db 321 GGGGTCTTCGGCCTTCCATCTCTAGGGTGCCGGCTCGGCTTAGGGCTCC 380

Qy 984 TCCGGGCTCACCACTCAGTACGGCGCCCGCCCTTCTCTGGGTCACGCTGGCAACCGGGCGTCCTG 1043
Db 381 TCCGCG-TCANCAATCAGTAACGAACCGCTTCTGGGTGCAAANGGGTCCTG 439

Qy 1044 TGCCTCTCCT 1054
Db   440 TGCCTNTCCT 450

; RESULT 14
US-09-759-130B-424
; Sequence 424, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MPI00-535OMNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24

; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
;
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 424
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-130B-424

Query Match   25.0%;   Score 354.8; DB 3; Length 1029;
Best Local Similarity   64.5%;   Pred. No. 3.4e-99;
Matches   546; Conservative   0; Mismatches 297; Indels   3; Gaps   1;

Qy   287 ACTGCCTTTTTACCCCCAGCCCCGGGCCATGCGCAGGCTTCAGCGTTCCACTGCTCATCGT 346
                 ||||| ||| || ||||| | |||||||||||| |||||||||| ||| ||| |||||
Db   21 ATTCCCCTTCTATGCTGGCCCCAAGCCAACCTTCCCGGATGGACACCACTTTGGCCAGCAT 80

Qy   347 TATTCTTAGTGTTTTTTGGCTCTAGCAGCAAGCTTCCTGCTCATCTTGCCCGGGATCCGTGG 406
                 | | |||| ||||| || |||| || | |||| ||||||||| ||| |||||
Db   81 CATCATGATCTTTCTGACTGCACTGGCCACGTTCATCGTCATCCTGCCTGGCATTCGGGG 140

Qy   407 CCACTCGCGCGCTGGTTTGGTTGGTTGGAGAGTTCTTCTCAGTCGTCATAGGCGCAGAAAT 466
                 || | | | || || ||| || |||||||||||| | ||||| | |||||||| ||
Db   141 AAAGACGAGGCGTGTTCTGGCTGCTTCGGGTGGTTGGACCAGCTTATTCATCGGGCTGCAAT 200

Qy   467 TGTGGCTGTGCACTTCAGTGCAGAATGGTTCGTGGGTACAGTGAACACCAACACATCCTA 526
                 | || | |||||||||| ||||| ||||||| || ||| ||| || |||| ||||
Db   201 CCTGGCTGTGAATTTCAGTTCTCAGTGGTCTGTGGGCCCAGGTCAGCACCAACACATCATA 260

Qy   527 CAAAGCCTTCAGCGCGAGGCGCCGCGTTACACGCCCGTGTCTCGGTCTCGTGTCTCGGTGG 586
                 |||||||| ||| |||| | |||||||||| ||||||| || ||| |||||| |||||
Db   261 CAAGGCCTTCAGTTCTGAGTGGATCACGCCCGTCTCGTTGGGGCCCAGGTCAGCACCAACACATCATA 320

Qy   587 CATTAATATTACACTCACAGGGACACCCCAGTCATCAGCGTGAACGAGACCATTGACTACAA 646
                 || |||| || ||||||| |||| ||||||||||||| |||| ||||| | ||| |||||
Db   321 AGTCAACATCACACTCACAGGGACCCCCGTCAGCAGGCGCTGAAATGAGACCATCAATTACAA 380

Qy   647 CGAGCAGTTCACCTGGCGCCTGGGTGTCGAAAGAGAATTACCGCCGCAGCTGAATGAGAGA 706
                 ||||||||||||||||| ||||| || ||| || || | |||| ||||||||| ||||
Db   381 CGAGGAGTTCACCTGGCGCCTGGGTGAGAACTATGCTGAGGAGTGTGCAAAAGGCTCTGGA 440

Qy   707 GAAGGGGGCTGCCGGACCCAGTGTCTCTACTCTGGGGAGAGTTCACACCGAGTAGCCCATG 766
                 | ||||||||| || | ||| ||| |||| |||| | |||| ||| | |||| ||| ||
Db   441 GAAGGGGGCTGCCAGACCCTGTGTTGTACCTAGCTGAGAAGTTCACTCAAGAAGCCCATG 500

Qy   767 CGGGCCTGTACCACCAGTACCACCTGGGCGGGGACACTAGCGCCTTCGGCGTATGGGTGGGC 826
                 ||||||| || ||||||| || |||||| ||| |||||||| ||| || |||| |||||
Db   501 TGGGCTGTGAACTTCAGTTCTCTGAAATTCAGTTCGTGTGGGGTCCAGGTCAGCAGCACTGTGGG 560

Qy   827 GTTCTCGCTTCTGGGCCTCTCCAACGGTGCTGGCGGCCGCCTCTACGGAGG 886
                 ||| | |||||||| || ||| |||| | || | | ||||| ||||||
Db   561 AATTCCTCTCGCTGTGGGCTGAGGCCAAATGATGGCCTCCATGCCGTGGTATATGGTGG 620

Qy   887 CCTGGCACTGCTGACCACCCGGAGCCTTCGGCGGCCTTCTTCG---GGGTCTTCGGCCCTTGGCCTC 943
                 ||||| | | | || | |||| | |||| || |||| ||||| |||| || || |||
Db   621 CTACATGCTATTGGCCACGGGCATCTTCCTGTTGGGCTTGGGCCTCTGTGTTTCTTCCATGGC 680

Qy   944 CATCTCTAGCCTGTGCCGCGCCTCGCCGCTAGCCTCTTCTCCGGCGCTTCTGTTCGTGCTCGCATACTCACCA 1003
                 ||| | | ||| ||| || |||||| |||| |||| |||||||||||| ||| ||||| | ||||| ||||
Db   681 CACATCACTCACCTCACCTTGTCCCTGACCTGTGCCTTAGGGTGCTGACACCTGAGAAGTTCACACCGAGTAGCCCTTG 740

; RESULT 15
US-10-042-431-54
; Sequence 54, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
;   APPLICANT: MCCARTHY, Sean A
;   APPLICANT: BARNES, Thomas M
;   APPLICANT: FRASER, Christopher C
;   APPLICANT: SHARP, John D
;   TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
;   TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
;   FILE REFERENCE: 10147-6U2
;   CURRENT APPLICATION NUMBER: US/10/042,431
;   CURRENT FILING DATE: 2001-10-25
;   PRIOR APPLICATION NUMBER: US 09/333,159
;   PRIOR FILING DATE: 1999-06-14
;   PRIOR APPLICATION NUMBER: US 09/578,063
;   PRIOR FILING DATE: 2000-05-24
;   NUMBER OF SEQ ID NOS: 79
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 54
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-431-54

Query Match   25.0%;   Score 354.8; DB 5; Length 1029;
Best Local Similarity   64.5%;   Pred. No. 3.4e-99;
Matches   546; Conservative   0; Mismatches 297; Indels   3; Gaps   1;

Qy   287 ACTGCCTTTTTACCCCCAGCCCCGGGCCATGCGCAGGCTTCAGCGTTCCACTGCTCATCGT 346
                 ||||| ||| || ||||| | |||||||||||| |||||||||| ||| ||| |||||
Db   21 ATTCCCCTTCTATGCTGGCCCCAAGCCAACCTTCCCGGATGGACACCACTTTGGCCAGCAT 80

Qy   347 TATTCTTAGTGTTTTTTGGCTCTAGCAGCAAGCTTCCTGCTCATCTTGCCCGGGATCCGTGG 406
                 | | |||| ||||| || |||| || | |||| ||||||||| ||| |||||
Db   81 CATCATGATCTTTCTGACTGCACTGGCCACGTTCATCGTCATCCTGCCTGGCATTCGGGG 140

Qy   407 CCACTCGCGCGCTGGTTTGGTTGGTTGGAGAGTTCTTCTCAGTCGTCATAGGCGCAGAAAT 466
                 || | | | || || ||| || |||||||||||| | ||||| | |||||||| ||
Db   141 AAAGACGAGGCGTGTTCTGGCTGCTTCGGGTGGTTGGACCAGCTTATTCATCGGGCTGCAAT 200

Qy   467 TGTGGCTGTGCACTTCAGTGCAGAATGGTTCGTGGGTACAGTGAACACCAACACATCCTA 526
                 | || | |||||||||| ||||| ||||||| || ||| ||| || |||| ||||
Db   201 CCTGGCTGTGAATTTCAGTTCTCAGTGGTCTGTGGGCCCAGGTCAGCACCAACACATCATA 260

Qy   527 CAAAGCCTTCAGCGCGAGGCGCCGCGTTACACGCCCGTGTCTCGGTCTCGTGTCTCGGTGG 586
                 |||||||| ||| |||| | |||||||||| ||||||| || ||| |||||| |||||
Db   261 CAAGGCCTTCAGTTCTGAGTGGATCACGCCCGTCTCGTTGGGGCCCAGGTCAGCACCAACACATCATA 320

Qy   587 CATTAATATTACACTCACAGGGACACCCCAGTCATCAGCGTGAACGAGACCATTGACTACAA 646
                 || |||| || ||||||| |||| ||||||||||||| |||| ||||| | ||| |||||
Db   321 AGTCAACATCACACTCACAGGGACCCCCGTCAGCAGGCGCTGAAATGAGACCATCAATTACAA 380

Qy   407 CCACTCGCGCGCTGGTTTGGTTGGTTGGAGAGTTCTTCTCAGTCGTCATAGGCGCAGAAAT 466
                 || | | | || || ||| || |||||||||||| | ||||| | |||||||| ||
Db   141 AAAGACGAGGCGTGTTCTGGCTGCTTCGGGTGGTTGGACCAGCTTATTCATCGGGCTGCAAT 200

Qy   467 TGTGGCTGTGCACTTCAGTGCAGAATGGTTCGTGGGTACAGTGAACACCAACACATCCTA 526
                 | || | |||||||||| ||||| ||||||| || ||| ||| || |||| ||||
Db   201 CCTGGCTGTGAATTTCAGTTCTCAGTGGTCTGTGGGCCCAGGTCAGCACCAACACATCATA 260

Qy   527 CAAAGCCTTCAGCGCGAGGCGCCGCGTTACACGCCCGTGTCTCGGTCTCGTGTCTCGGTGG 586
                 |||||||| ||| |||| | |||||||||| ||||||| || ||| |||||| |||||
Db   261 CAAGGCCTTCAGTTCTGAGTGGATCACGCCCGTCTCGTTGGGGCCCAGGTCAGCACCAACACATCATA 320

Qy   587 CATTAATATTACACTCACAGGGACACCCCAGTCATCAGCGTGAACGAGACCATTGACTACAA 646
                 || |||| || ||||||| |||| ||||||||||||| |||| ||||| | ||| |||||
Db   321 AGTCAACATCACACTCACAGGGACCCCCGTCAGCAGGCGCTGAAATGAGACCATCAATTACAA 380

Qy   647 CGAGCAGTTCACCTGGCGCCTGGGTGTCGAAAGAGAATTACCGCCGCAGCTGAATGAGAGA 706
                 ||||||||||||||||| ||||| || ||| || || | |||| ||||||||| ||||
Db   381 CGAGGAGTTCACCTGGCGCCTGGGTGAGAACTATGCTGAGGAGTGTGCAAAAGGCTCTGGA 440

Qy   707 GAAGGGGGCTGCCGGACCCAGTGTCTCTACTCTGGGGAGAGTTCACACCGAGTAGCCCTTG 766
                 | ||||||||| || | ||| ||| |||| |||| | |||| ||| | |||| ||| ||

```
Db   441 GAAGGGGGCTGCCAGACCCTGTGTTGTTGTACCTAGCTGAGAAGTTCACTCCAAAGAAGCCCATG  500
         |||||||||| |||| |||| || |||   ||    | ||||||||||||| ||  |  |||||| ||

Qy   767 CGGGCCTGTACCACCAGTACCACTGGGGGGACTACGCCTCGGGCCACGCTATGGGTGGGC  826
         |||||| ||| |||| |||

Db   501 TGGCCTATACCGCCAGTACCGCCTGGGCGGGACACTACACCTCAGCCATGCTATGGGTGGGC  560
         |||

Qy   827 GTTCTGCTTCTGGCCTCCTTCTCCAAACGTNGCTYGTCTCCCAAGGCCCGGCCCCCGCTCTTACGGAGG  886
         || |||| |||| |||||  |||||||||||  |||||||||| |||| |||

Db   561 AATTCCTCTGCTGGCTGCTGGGCCTGGCCCAAATGTGATGCTTCTCCATGCCTGTGCTGGGTATPATGTGGG  620
         ||||  ||| |||

Qy   887 CCTGGGCACTGCTGACCACCGGAGCCTTCGCGCTCTTCG---GGGTCTTCGGCCTTGGCCTC  943
         ||  |||||||  |||||||||||| ||  |||| |||     |||||||| |||||||| |

Db   621 CTACATGCTATTGGCACGGGGCATCTTCCAGCTGTTGGCTCTGCTTCTTCTTCCATGGC  680
         || |

Qy   944 CATCTCTTAGCGTGCCGGCTCTGCCCGCTCCGGCTCCTCCGGGCTTCACCACTCAGTTA  1003
         || |    |||| |||||||||||  |||||| ||  |||||| |||||

Db   681 CACATCACTCACTTCACCCTGTCCCTGCACCTGGGCGGTTCTGTTGGTGCATACTCACCA  740
         | |

Qy  1004 CGGCGCGCCGCCTTTCTGGGTCACGGCTGGCAACCGGCGTCCTGTNGCCTCTTCCTCCGGAGGGGC  1063
         |||||| ||  ||||||| ||||   ||||||| ||||||| |||| |||||||||| |||||||

Db   741 TGGGCCTGCCTTCTGGATCACATTGACCACACAGGGCTGCTGTGTGTGCTGCTCGGGGCCTGGC  800
         |

Qy  1064 CGTGGTGAGTCTCCAGTATGTTCGGCCCAGGCGCTCTTCGGCACCCTTCTGGACCAAAAGGGC  1123
         ||

Db   801 TATGGGCCGGTGGGCCCACAGGATGCAGCCTTCACAGGCTGAAAGGCTTTTCTTCAACCAGAGTGT  860
         |||

Qy  1124 CAAGGA  1129
         ||||

Db   861 GGATGA  866
```

OM nucleic - nucleic search, using sw model

```
                                GenCore version 5.1.6
                                Copyright (c) 1993 - 2006 Compugen Ltd.

Run on:                January 21, 2006, 22:44:31 ; Search time 320 Seconds
                                                     (without alignments)
                                                     3660.288 Million cell updates/sec

Title:          US-09-937-059-57
Perfect score:  1421
Sequence:       1 gaaccaggaaagtaacggct............ctctttttccttgtttctt 1421

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       6049916 seqs, 412136615 residues
Total number of hits satisfying chosen parameters:    1209832

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Published Applications NA_New:*
        1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
        2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
        3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
        4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
        5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
        6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
        7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
        8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
        9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
       10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
       11: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result                          Query
No.     Score   Match   Length  DB   ID                          Description
-----   -----   -----   ------  --  ------------------------     -----------
  1     44.8    3.2     2319     8  US-11-136-527-1934           Sequence 1934, Ap
  2     43      3.0     1006     7  US-10-821-234-422            Sequence 422, App
  3     43      3.0     1071     8  US-11-000-688-1164           Sequence 1164, Ap
c 4     42.6    3.0     88421    8  US-11-205-109-1              Sequence 1, Appli
c 5     41.8    2.9     88421    8  US-11-205-109-1              Sequence 1, Appli
c 6     41.6    2.9     1400     8  US-11-136-527-4440           Sequence 4440, Ap
c 7     41.6    2.9     2305     8  US-11-136-527-344            Sequence 344, App
c 8     41.6    2.9     12309    7  US-10-995-561-13450          Sequence 13450, A
c 9     41.6    2.9     95832    7  US-10-995-561-13273          Sequence 13273, A
c 10    41.4    2.9     173602   8  US-11-121-086-25             Sequence 25, Appl
c 11    41      2.9     2432     7  US-10-750-185-29908          Sequence 29908, A
c 12    41      2.9     2432     7  US-10-750-623-29908          Sequence 29908, A
c 13    41      2.9     3886     8  US-11-087-100-37             Sequence 37, Appl
c 14    41      2.9     3886     8  US-11-087-084-37             Sequence 37, Appl
c 15    41      2.9     3886     8  US-11-087-085-37             Sequence 37, Appl
c 16    40.8    2.9     2862     8  US-11-037-243-7              Sequence 7, Appli
c 17    40.4    2.8     201      7  US-11-075-400-17             Sequence 17, Appl
c 18    40.4    2.8     11438    8  US-10-821-234-224            Sequence 224, App
c 19    40      2.8     2797     8  US-11-075-400-17             Sequence 17, Appl
c 20    40      2.8     11612    7  US-10-995-561-499            Sequence 499, App
c 21    39.8    2.8     717      7  US-10-883-512-113            Sequence 113, App
c 22    39.8    2.8     895      7  US-10-883-512-109            Sequence 109, App
c 23    39.8    2.8     895      7  US-10-883-512-112            Sequence 112, App
c 24    39.8    2.8     955      7  US-10-883-512-110            Sequence 110, App
c 25    39.8    2.8     34875    7  US-10-775-169-316            Sequence 316, App
c 26    39.8    2.8     164810   8  US-11-121-086-4              Sequence 4, Appli
c 27    39.8    2.8     172543   8  US-11-121-086-6              Sequence 6, Appli
c 28    39.2    2.8     2133     8  US-11-143-980-13             Sequence 13, Appl
c 29    39.2    2.8     116856   8  US-11-143-980-1              Sequence 1, Appli
c 30    38.8    2.8     10800    8  US-11-019-711-5              Sequence 5, Appli
c 31    38.8    2.7     10809    8  US-11-019-711-1              Sequence 1, Appli
c 32    38.8    2.7     1494     8  US-11-143-980-26             Sequence 26, Appl
c 33    38.4    2.7     157224   8  US-11-112-908-51             Sequence 51, Appl
c 34    38.4    2.7     161726   8  US-11-112-908-48             Sequence 48, Appl
c 35    38.4    2.7     161726   8  US-11-112-908-52             Sequence 52, Appl
c 36    38.2    2.7     170189   8  US-11-112-908-50             Sequence 50, Appl
c 37    38.2    2.7     850      8  US-10-883-512-114            Sequence 114, App
c 38    38.2    2.7     3200     7  US-10-645-441-13             Sequence 13, Appl
c 39    38      2.7     1897     7  US-10-775-169-236            Sequence 236, App
c 40    38      2.7     1110     8  US-11-112-908-14             Sequence 14, Appl
c 41    37.8    2.7     8651     7  US-10-432-483-48             Sequence 48, Appl
c 42    37.8    2.7     963      7  US-10-858-730-182            Sequence 182, App
c 43    37.8    2.7     1725     7  US-10-821-234-772            Sequence 772, App
c 44    37.8    2.7     1883     8  US-11-012-668-9              Sequence 9, Appl
c 45    37.2    2.6     1527     8  US-11-075-185-56             Sequence 56, Appl
```

ALIGNMENTS

RESULT 1
US-11-136-527-1934
; Sequence 1934, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1934
; LENGTH: 2319
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-1934

  Query Match   3.2%; Score 44.8; DB 8; Length 2319;
  Best Local Similarity   51.5%;
  Matches 103; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

Qy   857 GCTCTCCACGCGCGGCCCCGCGCTCTACGGAGGGCTGGCACTGCTGACCACGGAGCCTTCGC 916
              ||||| || | | || || ||| | || | || ||| ||| ||| || || || || |
Db   776 GCACGCGCGTGGTGTCTCTGCCGCCGACGGGGCGCCTCTTCGTCGTTCATGTTTCGCGTGGGGGGGGACT 835

Qy   917 GCTCTCTCCGGGGTGTCTTTGCCCTTGGCCTCCATCTTCTAGCGTGCGGCGTGCGGCTCTGCGCGCCCGCCTCCCGCT 976
              | ||| | | || | | | || | || | | || | | | | | || | | | |
Db   836 GCGATCCTCGCACATCGTCTCGAGGCCTCCATCCGCGCGCCAAGCTCATCCGCTCGCTCGCGCGCCAGAC 895

Qy   977 AGGCTCCTTCGCGGCTCACCACTCAGTACGGCGCGCCGCGCTTCTCGGGTCACGGCTCGGCAACCGG 1036
              | |||| ||| || || | || || | || ||| | | || | | | ||| | ||
Db   896 GCTCGAGGGCGGAGTTCATCCCTCTGCACCAGACCACTTCACGGTGGGCGTTCGACACGGG 955

Qy   1037 CGTCCTGTGCCTCTTCCTCG 1056
              | | ||| |||| || ||
Db   956 GGACGACCGCCTCTTCCTCG 975

RESULT 2
US-10-821-234-422
; Sequence 422, Application US/10821234

```
; Publication No. US20050255114A1
; GENERAL INFORMATION:
;   APPLICANT:
;     APPLICANT: Labat, Ivan
;     APPLICANT: Stache-Crain, Birgit
;     APPLICANT: Andarmani, Susan
;     APPLICANT: Tang, Y. Tom
;   TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;   FILE REFERENCE: 821A
;   CURRENT APPLICATION NUMBER: US/10/821,234
;   CURRENT FILING DATE: 2004-04-07
;   PRIOR APPLICATION NUMBER: US 60/462,047
;   PRIOR FILING DATE: 2003-04-07
;   NUMBER OF SEQ ID NOS: 1704
;   SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 422
;   LENGTH: 1006
;   TYPE: DNA
;   ORGANISM: Homo sapiens

US-10-821-234-422

Query Match      3.0%;  Score 43;  DB 7;  Length 1006;
Best Local Similarity 48.9%;  Pred. No. 0.18;
Matches 115;  Conservative  0;  Mismatches 120;  Indels  0;  Gaps  0;

Qy  722 CCCAGTGTCTACCTGGGCGGGAGGAGAAGTTCACACCGAGTAGCCCTTGGCGGCGGCCTGTACCACCA 781
            |||| |||| |||||||| || |||||||    ||||  |||||| || ||||||| ||| | |||||
Db  295 CCTGCTGCTACAGCTGTCGGAGGAAGGCCACACACTGGGGCCCTTGCGTGCCGCCCCTGCCCTG 354

Qy  782 GTACCACCTGGCGGGGACACTACCGCCTCGGCGCAGGCTATGGGTGGGCGGTTCTGCTTCTTCGGCT 841
            |||||||| ||||||| || ||||| ||||||||||| ||| || ||| ||| ||||| || |||||
Db  355 GCAGCGCGTGGACCGCGACGTGGCACCGGGAACTTCTTCTGCGACGTGGGCCGGCTGGGGCAT 414

Qy  842 CCTCTCCAACGTGCTGTCTTCCAGCCCGGCGCCCCGCCTCTACGGAGGCCTGCCACTGCTGAC 901
            |||||| |||||||||| ||| |||||||||||||| |||| |||||||||||| |||||||
Db  415 AGTCAACCACGCCGGGCGCCGCCCCGGACGCGCCTGCTGCTTTGCCAGTCCCAGTGCTGGA 474

Qy  902 CACCGGAGCCTTCGCGGCGTCTTCGGGGTGTTTGCGGGGCCTCCAATCTCTTAGCGGTG 956
            ||||| || ||||| ||||||||||||| ||||||||| ||| |||||| ||| |||||
Db  475 CCGCGCCACCTGCAACCGGCGCACGCACCACGAGGGGCCCATCACCGAGCGCTTG 529
```

```
RESULT 3
US-11-000-688-1164
; Sequence 1164, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
;   APPLICANT:
;     APPLICANT: BERTUCCI, Francois
;     APPLICANT: HOULGATTE, Remi
;     APPLICANT: BIRNBAUM, Daniel
;   TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
;   FILE REFERENCE: 1423-R-03
;   CURRENT APPLICATION NUMBER: US/11/000,688
;   CURRENT FILING DATE: 2004-12-01
;   PRIOR APPLICATION NUMBER: US 60/525,987
;   PRIOR FILING DATE: 2003-12-01
;   NUMBER OF SEQ ID NOS: 1596
;   SOFTWARE: Patentin version 3.2
; SEQ ID NO 1164
;   LENGTH: 1071
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial sequences:primer
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (1)..(1071)
;     OTHER INFORMATION: d component of complement of complement (adipsin) (DF)
;     OTHER INFORMATION: gene.

US-11-000-688-1164

Query Match      3.0%;  Score 43;  DB 8;  Length 1071;
Best Local Similarity 48.9%;  Pred. No. 0.18;
Matches 115;  Conservative  0;  Mismatches 120;  Indels  0;  Gaps  0;
```

```
RESULT 4
US-11-205-109-1/c
; Sequence 1, Application US/11205109
; Publication No. US20050287641A1
; GENERAL INFORMATION:
;   APPLICANT:
;     APPLICANT: Farnet, Chris
;     APPLICANT: Zazopoulos, Emmanuel
;     APPLICANT: Staffa, Alfredo
;   TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
;   FILE REFERENCE: 3002-2US
;   CURRENT APPLICATION NUMBER: US/11/205,109
;   CURRENT FILING DATE: 2005-08-17
;   PRIOR APPLICATION NUMBER: US/09/976,059
;   PRIOR FILING DATE: 2001-10-15
;   PRIOR APPLICATION NUMBER: US 60/239,924
;   PRIOR FILING DATE: 2000-10-13
;   NUMBER OF SEQ ID NOS: 46
;   SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
;   LENGTH: 88421
;   TYPE: DNA
;   ORGANISM: Actinoplanes sp.
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (2077)..(3078)
;     OTHER INFORMATION: ORF 1; positive strandedness
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (3118)..(4032)
;     OTHER INFORMATION: ORF 2; positive strandedness
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (4038)..(5048)
;     OTHER INFORMATION: ORF 3; positive strandedness
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (6665)..(5814)
;     OTHER INFORMATION: ORF 4; negative strandedness
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (7703)..(6693)
;     OTHER INFORMATION: ORF 5; negative strandedness
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (9464)..(8130)
;     OTHER INFORMATION: ORF 6; negative strandedness
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (9691)..(10761)
;     OTHER INFORMATION: ORF 7; positive strandedness
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (12751)..(10829)
;     OTHER INFORMATION: ORF 8; negative strandedness
;   FEATURE:
```

```
Qy  722 CCCAGTGTCTACCTGGGCGGGAGGAGAAGTTCACACCGAGTAGCCCTTGGCGGCGGCCTGTACCACCA 781
            |||| |||| |||||||| || |||||||    ||||  |||||| || ||||||| ||| | |||||
Db  324 CCTGCTGCTACAGCTGTCGGAGGAAGGCCACACACTGGGGCCCTTGCGTGCCGCCCCTG 383

Qy  782 GTACCACCTGGCGGGGACACTACCGCCTCGGCGCAGGCTATGGGTGGGCGGTTCTGCTTCTTCGGCT 841
            |||||||| ||||||| || ||||| ||||||||||| ||| || ||| ||| ||||| || |||||
Db  384 GCAGCGCGTGGACCGCGACGTGGCACCGGGAACTTCTTCTGCGACGTGGGCCGGCTGGGGCAT 443

Qy  842 CCTCTCCAACGTGCTGTCTTCCAGCCCGGCGCCCCGCCTCTACGGAGGCCTGCCACTGCTGAC 901
            |||||| |||||||||| ||| |||||||||||||| |||| |||||||||||| |||||||
Db  444 AGTCAACCACGCCGGGCGCCGCCCCGGACGCGCCTGCTGCTTTGCCAGTCCCAGTGCTGGA 503

Qy  902 CACCGGAGCCTTCGCGGCGTCTTCGGGGTGTTTGCGGGGCCTCCAATCTCTTAGCGGTG 956
            ||||| || ||||| ||||||||||||| ||||||||| ||| |||||| ||| |||||
Db  504 CCGCGCCACCTGCAACCGGCGCACGCACCACGAGGGGCCCATCACCGAGCGCTTG 558
```

NAME/KEY: misc_feature
OTHER INFORMATION: ORF 9; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13617)..(12802)
OTHER INFORMATION: ORF 10; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15203)..(13614)
OTHER INFORMATION: ORF 11; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15591)..(15863)
OTHER INFORMATION: ORF 12; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15880)..(19035)
OTHER INFORMATION: ORF 13; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (19032)..(39713)
OTHER INFORMATION: ORF 14; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (39713)..(65800)
OTHER INFORMATION: ORF 15; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65826)..(66530)
OTHER INFORMATION: ORF 16; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (66546)..(67370)
OTHER INFORMATION: ORF 17; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (67384)..(70059)
OTHER INFORMATION: ORF 18; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (70099)..(70662)
OTHER INFORMATION: ORF 19; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (70659)..(71906)
OTHER INFORMATION: ORF 20; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (73459)..(71964)
OTHER INFORMATION: ORF 21; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (74216)..(73563)
OTHER INFORMATION: ORF 22; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (75424)..(74211)
OTHER INFORMATION: ORF 23; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (75555)..(76464)
OTHER INFORMATION: ORF 24; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (78110)..(76449)
OTHER INFORMATION: ORF 25; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (79864)..(78107)
OTHER INFORMATION: ORF 26; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (81624)..(79861)
OTHER INFORMATION: ORF 26; negative strandedness
FEATURE:
NAME/KEY: misc_feature

LOCATION: (81909)..(81682)
OTHER INFORMATION: ORF 27; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (82346)..(82062)
OTHER INFORMATION: ORF 28; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (82587)..(84446)
OTHER INFORMATION: ORF 29; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84481)..(85548)
OTHER INFORMATION: ORF 30; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (85556)..(86845)
OTHER INFORMATION: ORF 31; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (87372)..(86803)
OTHER INFORMATION: ORF 32; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (87494)..(88420)
OTHER INFORMATION: ORF 33; positive strandedness; N-terminus only

US-11-205-109-1

Query Match   3.0%; Score 42.6; DB 8; Length 88421;
Best Local Similarity   46.5%; Pred. No. 1.4;
Matches 138; Conservative   0; Mismatches 159; Indels   0; Gaps   0;

```
Qy     839 GCTCTCTCCAACGTGCTCGTCTCCACGCGCCGGCCCCGGCTCTTACGGAGGCCTGGCACTGCT 898
           |||||||  |||| | ||| || |||||||  ||||| ||  | || ||   | ||||| ||||
Db   13270 GGTGCGCCTGCGCGGGTCCTGATCGCGACCCTGGCCCCGACCGCCGCCCCGGCGCGCGCGCA 13211

Qy     899 GACCACCGGAGCCCTTCGCGCGTCTTCGGGGTCTTTCGCCTTGGCCTTCCATCTCTAGCCTGCC 958
           |||  |||  |||  ||   |||||| |||||||| |||||||| |||||||||  | ||||||
Db   13210 GAACCAGGCCCGACCATCGACCTTGCTGGAGGAGGGGGGTCACGCTGCCCGCCCGCGTACGGAGGC 13151

Qy     959 GCTCTTGCCCCGCGCCGCCTAGGCTCCTTCCCGCGGCTCACCCACTCAGTTACGGCGCGGCGCCTTCTG 1018
           ||| | |||| ||| || |||  || |   |||||||| || ||||| || ||||| || || | |
Db   13150 CGCGACCGCGCGGTCTTCAAGATAGTTCTCCCCGGCGACCGCTCAAACGACGACGTGGCCCGTGCG 13091

Qy    1019 GGTCACGCTGGCAACCGGCGGTCCTGTGCCTCCTTCCTCCTGGAGGGGGCCCTGGTGGAGGTCTCCA 1078
           |||  || |   | || | ||||  || || |||| |  || || || | |||| ||| || ||||
Db   13090 CGAGTGGCCTCGACATCTTCGAGCTCTTCCGGCACGGACGCGGTCTTCGGCCCCGGGGCAGGCCTG 13031

Qy    1079 GTATGTTTCGGCCCCACACGGCTCTTTCGCCGACCCTTTCTGACCAAAAGCGCCAAAGGACTGCAG 1135
           |    |  ||| |||| | |||| || |||| |  | ||| ||| || |||| || | | ||||
Db   13030 GGCCGAGCTGACCGGGTGACCGGCGCGCGCCCCCCTGCGTCGGGTCACCGCGCCCCTGCGCG 12974
```

RESULT 5
US-11-205-109-1

Sequence 1, Application US/11205109
Publication No. US20050287641A1
GENERAL INFORMATION:
APPLICANT: Farner, Chris
APPLICANT: Zazopoulos, Emmanuel
APPLICANT: Staffa, Alfredo
TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
FILE REFERENCE: 3002-2US
CURRENT APPLICATION NUMBER: US/11/205,109
CURRENT FILING DATE: 2005-08-17
PRIOR APPLICATION NUMBER: US/09/976,059
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US 60/239,924
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 88421

TYPE: DNA
ORGANISM: Actinoplanes sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2077)..(3078)
OTHER INFORMATION: ORF 1; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3118)..(4032)
OTHER INFORMATION: ORF 2; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4038)..(5048)
OTHER INFORMATION: ORF 3; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (6665)..(5814)
OTHER INFORMATION: ORF 4; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (7703)..(6693)
OTHER INFORMATION: ORF 5; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9464)..(8130)
OTHER INFORMATION: ORF 6; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9691)..(10761)
OTHER INFORMATION: ORF 7; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (12751)..(10829)
OTHER INFORMATION: ORF 8; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13617)..(12802)
OTHER INFORMATION: ORF 9; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15203)..(13614)
OTHER INFORMATION: ORF 10; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15591)..(15863)
OTHER INFORMATION: ORF 11; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15880)..(19035)
OTHER INFORMATION: ORF 12; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (19032)..(19713)
OTHER INFORMATION: ORF 13; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (39713)..(65800)
OTHER INFORMATION: ORF 14; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65826)..(66530)
OTHER INFORMATION: ORF 15; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65546)..(67370)
OTHER INFORMATION: ORF 16; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (67384)..(70059)
OTHER INFORMATION: ORF 17; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (70099)..(70662)
OTHER INFORMATION: ORF 18; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (70659)..(71906)
OTHER INFORMATION: ORF 19; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (73439)..(71964)
OTHER INFORMATION: ORF 20; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (74216)..(73563)
OTHER INFORMATION: ORF 21; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (75454)..(74213)
OTHER INFORMATION: ORF 22; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (75535)..(76464)
OTHER INFORMATION: ORF 23; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (78110)..(76449)
OTHER INFORMATION: ORF 24; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (79864)..(78107)
OTHER INFORMATION: ORF 25; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (81624)..(79861)
OTHER INFORMATION: ORF 26; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (81909)..(81682)
OTHER INFORMATION: ORF 27; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (82346)..(82062)
OTHER INFORMATION: ORF 28; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (82587)..(84446)
OTHER INFORMATION: ORF 29; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84481)..(85548)
OTHER INFORMATION: ORF 30; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (85556)..(86845)
OTHER INFORMATION: ORF 31; positive strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (87372)..(86803)
OTHER INFORMATION: ORF 32; negative strandedness
FEATURE:
NAME/KEY: misc_feature
LOCATION: (87494)..(88420)
OTHER INFORMATION: ORF 33; positive strandedness; N-terminus only

US-11-205-109-1

```
Query Match       2.9%;   Score 41.8;  DB 8;  Length 88421;
Best Local Similarity   50.8%;   Pred. No. 2.2;
Matches  100;  Conservative    0;  Mismatches  97;  Indels   0;  Gaps   0;

Qy     805 CCTCGGCCACCGCTATCGGGTGGCGTTCTGCTTCTGGCTCCTTCCCAACGTGCTGCTCTCCA 864
           |||| ||| |   | ||| ||||||||||||||||| |||||||  |||| || ||  |||
Db   73052 CGTCGACGCGCCCAGCCGCGCGCGGTTCTCCAGCAGGATCTGGTCGAAACCGCGGCGGCTCCA 73111

Qy     865 CGCCGGCCCCGCGCTTACGGAGGCGCTGGCACTGCTGGACCACCGGAGACCTTCGGCGCTCTTCG 924
           ||| ||||||||| ||| |||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 73112 CCTGGAACGCGTGCGAGGTCGGCCCGGGGGCGAGGTGGCGAAGGTGAACGTCC 73171
         |||||| |||| |||| |||||||| |||
Qy   925 GGGTCTTCGCCTTGGCCTTCCATCTCTACGGTGCCGGCTTCGCCCGCTCCGCCCCTAGGCTCCT 984
Db 73172 ACGGCTCGGTACTGGTGCCCACTTGACGTGCGCCCGTGCTTGCGCATGAACCGGCCT 73231
         |||| |||||| |||
Qy   985 CCGGCGCTCACCACTCAG 1001
Db 73232 TGGCGGATCTCGTCGCCG 73248
```

RESULT 6
US-11-136-527-4440/c
; Sequence 4440, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 4440
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-4440

Query Match 2.9%; Score 41.6; DB 8; Length 1400;
Best Local Similarity 35.3%; Pred. No. 0.47;
Matches 77; Conservative 42; Mismatches 99; Indels 0; Gaps 0;

```
Qy  845 CTCCAACGTGCTGCTCTCCAACGCGCCCGCTCTACGGAGGCTGGGCACTGCTGACCAC 904
Db  796 CTCRATCTKKCKTVKKAGCAGCTKGCTCMTCSYCAGCASTYTCSTCTKKTKCTKCTTCCT 737
Qy  905 CGGAGCCTTCGCGCGCTCTTCGGCGTCTTGGCCTTCCATCTCTAGCGTGCCGCCTCTG 964
Db  736 CTKCTRAYTCKKKGGTSTCTTSCKCTTCWYYCTCWCSDCCKTCTYCTKCTTKGSCMTSTT 677
Qy  965 CCCCGCTCCGCCCTAGGCTCCTCACCACTCAGTACGGCGCCGCCGCCTTCTGGGTCAC 1024
Db  676 CWGRCTCATYYTTGDCAGCTYSTRYCKCCTCWGCAYCTTCCTCTKYCTCCCCKSCTCTT 617
Qy 1025 GCTGGCAACGGGGCGTCCTGTGCCTCTTCCTCGGGAGGGG 1062
Db  616 TCTSCKTCYCTCYTCTCWTCTKCTTCCTTCARAGGGG 579
```

RESULT 7
US-11-136-527-344/c
; Sequence 344, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 344
; LENGTH: 2305
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-344

Query Match 2.9%; Score 41.6; DB 8; Length 2305;
Best Local Similarity 35.3%; Pred. No. 0.57;
Matches 77; Conservative 42; Mismatches 99; Indels 0; Gaps 0;

```
Qy  845 CTCCAACGTGCTGCTCTCCAACGCGCCCGCTCTACGGAGGCTGGGCACTGCTGACCAC 904
Db 1701 CTCRATCTKKCKTVKKAGCAGCTKGCTCMTCSYCAGCASTYTCSTCTKTCTKCTTCCT 1642
Qy  905 CGGAGCCTTCGCGCGCTCTTCGGCGTCTTGGCCTTCCATCTCTAGCGTGCCGCCTCTG 964
Db 1641 CTKCTRAYTCKKKGGTSTCTTSCKCTTCWYYCTCWCSDCCKTCTYCTKCTTKGSCMTSTT 1582
Qy  965 CCCCGCTCCGCCCTAGGCTCCTCACCACTCAGTACGGCGCCGCCGCCTTCTGGGTCAC 1024
Db 1581 CWGRCTCATYYTTGDCAGCTYSTRYCKCCTCWGCAYCTTCCTCTKYCTCCCCKSCTCTT 1522
Qy 1025 GCTGGCAACGGGGCGTCCTGTGCCTCTTCCTCGGGAGGGG 1062
Db 1521 TCTSCKTCYCTCYTCTCWTCTKCTTCCTTCARAGGGG 1484
```

RESULT 8
US-10-995-561-13450
; Sequence 13450, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13450
; LENGTH: 12309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13450

Query Match 2.9%; Score 41.6; DB 7; Length 12309;
Best Local Similarity 49.1%; Pred. No. 1.1;
Matches 107; Conservative 1; Mismatches 110; Indels 0; Gaps 0;

```
Qy  907 GAGCCTTCGCGCGCTCTTCGGCGGTTCTTGGCCTTCGGCTTGGCCTCCATCTCTAGCGTGCGCGCTCTGCC 966
Db 6026 GGGTCCCCGGCGCGCTCGGGCAGCGTGGGCGCCCGATGCGGGTGCCCGCGCGGG 6085
Qy  967 CGCTCCGCCCTAGGCTCCTCCCGGCTCACCACTCAGTACGGCGCCGCCTTCTGGGTCACGG 1026
Db 6086 CCCTCCCGGGGGAGCGGCGCCGCCTGCCGCCGCGCGCCCCCCGGCCGAGCCCGTCG 6145
Qy 1027 TGGCAACGGCGGCGTCCTGTGCCTCTTCCTCGGAGGGGCGCCGGTGGTGTGAGTCTTCCAGTTC 1086
Db 6146 CGCTCCATCGTVGCCCCGGCGCCTCCGGGCCGGCCGCCTTTAATAGGCCC 6205
Qy 1087 GGGCCCAGCGCTTCTTCGGCACCCTTCTGGACCAAAGCGGC 1124
Db 6206 CGCGCGCGCCCCCCGCYCGCCCCCCCCCCMCGCGCGCC 6243
```

RESULT 9
US-10-995-561-13273/c
; Sequence 13273, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559

; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 13273
; LENGTH: 95832
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13273

Query Match 49.1%; Score 41.6; DB 7; Length 95832;
Best Local Similarity 2.9%; Pred. No. 2.6;
Matches 107; Conservative 1; Mismatches 110; Indels 0; Gaps 0;

Qy 907 GAGCCTTCCGCGCTCTTCGGGGGTCTTTCGCCTTGGCGTGCCGCTCTGCC 966
Db 6145 GGGTCCCCGGGGCGGCCTCGGCAGCGTGGCTGGCGCGGGCCCCCGTTCCCCGCGGGG 6086

Qy 967 CGCTCGCGCCTAGGCGTCCTCCGCGGCGTCACCACTCAGTACGGCCGCCCTTCTGGGTCACGC 966
Db 6085 CCCTCCCGGGGGAGGGGCGGCGGCCGCGCCCTTGCGCCCCCCCGCGCAGCCGGTCG 6026

Qy 1027 TGGCAACCGGGGGGTCCTGTGTGCCTCTTCTCCTCGGGAGGGGGCCGTGGGTGAGTCTCCAGTTATGTTC 1086
Db 6025 CGCTCCACGTGGCCCCCCGGCCCGCGCTCTGGGGGCCCCSGCGCGCCGCCTTTAATAGGGCCC 5966

Qy 1087 GGCCCAGCGCCTCTTCGCACCCTTCTGGACCAAAGCGCC 1124
Db 5965 CGGCGGCCCCCGCCCCCGCCGGCYCGGCCCCTCCCCGMCGCGGCC 5928

RESULT 10
US-11-121-086-25/c
Sequence 25, Application US/11121086
Publication No. US200502664591A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138.6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin version 3.3
;
; SEQ ID NO 25
; LENGTH: 173602
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-25

Query Match 48.9%; Score 41.4; DB 8; Length 173602;
Best Local Similarity 0; Pred. No. 3.7;
Matches 111; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

Qy 829 TCTGCTTCTGGCTTCCTCTCCAACGTGCTGCTCTTCCACGCGGCCCCGGCTCTACGGAGGCC 888
Db 51356 TCGTCTTTCTTTCTTTCCTCTTCCTTCGGGCGCTTTCCCGCTCCGCGTCCTTCTCCATCT 51297

Qy 889 TGGCACTCGTGACCACCGGAGCCTTCGGCGTCTTCGCGTTGCGGGGTCTTCGCCTTGGCGTCCAGCT 948
Db 51296 TCCTGCTTCTCCTCCCCTTACTTTCTTTCTTCCCCTTCCCCTTCCCTTTCCCATCT 51237

Qy 949 CTAGGCGTGCCGCTCTCGGCCCTTCCGCCTTAGGCTCCTCCGCGCTTCACCACTCAGTACGGCG 1008
Db 51236 TCCTCGTCTCTCGTTCCCTTTCCCCTTCTCTCCTCTCTCCTCCTCCTCTCCCCT 51177

Qy 1009 CCGCCTTCTGGGGTCACGCTGGGCAACCGGGGGGTCCTGTGGCCTCTTCCTC 1055
Db 51176 TCCCTTTCCCTTTCCCTTCTCTCCTCTCTCCTCTTCCTTCCTTC 51130

;
; CURRENT APPLICATION NUMBER: US/10/750,185
RESULT 11
US-10-750-185-29908
Sequence 29908, Application US/10750185
Publication No. US200502660603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DeNISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MMI1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 29908
; LENGTH: 2432
; TYPE: DNA
; ORGANISM: Bovine 1986688091940
US-10-750-185-29908

Query Match 47.2%; Score 41; DB 7; Length 2432;
Best Local Similarity 0; Pred. No. 0.84;
Matches 125; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 788 CCTGGGCGGGGACACTACGGCCTCGGCCCACGCGTATGGGTGGGCGGTTCTTGCTTCGGCGTCCTCTC 847
Db 2073 CCTGGCCGGTGGCGGGGGGCATCGCGCGCCGCCACGCGTTCGTTCGTGGTTGGTGGTTCACCTTGGCGCTACGA 2132

Qy 848 CAACGTGCTGCTTCCACGCGGGCCCCGCGGCCGTTCACGGAGGCGCTGGGCGACTGCTGCTGACCACCGG 907
Db 2133 CGACACGCCCATCGTCAAGGCGCTCGGGCGCCGGGGAGCTGACGTAGGTGCTCTTGCGCGGGCAT 2192

Qy 908 AGCCTTCGGCGCTCTTCGGGGGTCTTCGCCTTGGCGTCCATCTCTAGGCGTGGCGGCGTCTGCCC 967
Db 2253 CTTCCTGTGCTACGGCCACCACCTTCCCTCACTCGCCGGAGGCCGACCTGGGGCCGCACTTGCTC 2252

Qy 968 GCTCGCGCCTAGGCGTCCTCCGCGGCTCACACACTCAGTACGCGGCGCGGGCCCTTCTGGGGTCACGCT 1027
Db 2253 TCTCGCGCCGGATCTTCCTGGGGCTCGGGCATGAGGCATCAGCTATGGGGCCCGCTCCACCAA 2312

Qy 1028 GGGCAACCGGGCGTCCTTGTGGCCTCTTC 1052
Db 2313 GACCAACCCCATCTACCGGCATCTTC 2337

RESULT 12
US-10-623-29908
Sequence 29908, Application US/10750623
Publication No. US2005028757531A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DeNISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MMI1100-1
CURRENT APPLICATION NUMBER: US/10/750,623
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 29908
; LENGTH: 2432

; TYPE: DNA
; ORGANISM: Bovine   1986688091914O
US-10-750-623-29908

Query Match   Best Local Similarity   47.2%;   Score 41; DB 7; Length 2432;
Matches 125; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

```
Qy    788 CCTGGCGGGACACTACCCTCGGCCAGCTATGGGTGGGTTCTGCTTCTGGCTCCTCTC 847
          |||||||||  ||||||  ||  ||| ||||  |||| |||||||||||||| ||||
Db   2073 CCTGGGCCGTGGGCGGGCATCGGCCCGCACGCCGGTGTGGTGCACCTTCGTGCGGCTACAA 2132

Qy    848 CAACGTGCTGCTCTCCACGCCCGGCCCCGCCTCTACGGAGGCCACTGCTGCTGACCACCGG 907
          |||||||   |||  |||| |||||||| || |||||||||  ||||||||||||||||
Db   2133 CGACACGCCCATCGTCAAGGCCTTCGGGCCGGACGCTGAGCTACGTGCTCTGCGGGGCAT 2192

Qy    908 AGCCTTCGCGCGCTTCTTCGGGGTTCTTCGCCTTGGGCCTCCATCTCTAGGCGTGCCGCCC 967
          ||||||||| ||||||||  |||||  || ||  ||||||| |||| ||||| ||| ||||
Db   2253 CTTCCGTGTCGTAGGCGCACCACCTTCCTCATTGATCGCCCGGAGCCCGGACCTTCGTCACCGCT 2252

Qy    968 GCTCCGCCTAGGCGTTCCTCCGGCGCTTCACCACTGCAGTACGGCGCGGCCTTTCTGGGTTCACCGCT 1027
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2253 TCTGCGCGCGGGATCTTCCTGGGGGTTTCGGCGATGAGCATCAGCTATGCGGGCGCCCTGCTTCACCAA 2312

Qy   1028 GGCAACCGCGGCGTCCTGTGTCCCTCTTC 1052
          ||||||||||||||||||||||||||||
Db   2313 GACCAACCGCATCTACCGGATCTTC 2337
```

RESULT 13
US-11-087-100-37/c
; Sequence 37, Application US/11087100
; Publication No. US20050266440A1
; GENERAL INFORMATION:
; APPLICANT: Metz, James
; APPLICANT: Barclay, William
; APPLICANT: Flatt, James
; APPLICANT: Kuner, Jerry
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase
; FILE REFERENCE: 2997-29
; CURRENT APPLICATION NUMBER: US/11/087,100
; CURRENT FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: 09/231,899
; PRIOR APPLICATION NUMBER: 1999-01-14
; PRIOR FILING DATE: 60/284,066
; PRIOR APPLICATION NUMBER: 2001-04-16
; PRIOR FILING DATE: 60/298,796
; PRIOR APPLICATION NUMBER: 2001-06-15
; PRIOR FILING DATE: 60/323,269
; PRIOR APPLICATION NUMBER: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SEQ ID NO 37
; LENGTH: 3886
; TYPE: DNA
; ORGANISM: Schizochytrium sp.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2115)..(2115)
; OTHER INFORMATION: n = a, c, g, or t
US-11-087-100-37

Query Match   Best Local Similarity   52.0%;   Score 41; DB 8; Length 3886;
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```
Qy    702 CTGGAGAAGGGGCGTCCGGACCCAGTGCTTCTACCTGGGCGGAGAAGTTCACACCGAGTAGC 761
          |||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3850 CTGCTGCTGCTGCTGCTGCTGCTGGGCGGAGAAGTTCACCAGTAGC 3791

Qy    762 CCTTGCGGCGGCCTGTACCACCAGTACCACCTGGGCGGGACACTACCGCCTTCGGCCACGGTATGG 821
          |||  ||   ||||||||||  |||||||||||||||||||||
Db   3850 CCTTGCGGCGGCCTGTACCACCAGTACCACCTGGGCGGGACACTACCGCCTTCGGCCACGGTATGG 821
```

RESULT 14
US-11-087-084-37/c
; Sequence 37, Application US/11087084
; Publication No. US20050273883A1
; GENERAL INFORMATION:
; APPLICANT: Metz, James
; APPLICANT: Barclay, William
; APPLICANT: Flatt, James
; APPLICANT: Kuner, Jerry
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase
; FILE REFERENCE: 2997-29
; CURRENT APPLICATION NUMBER: US/11/087,084
; CURRENT FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: 09/231,899
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: 60/284,066
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/298,796
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/323,269
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SEQ ID NO 37
; LENGTH: 3886
; TYPE: DNA
; ORGANISM: Schizochytrium sp.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2115)..(2115)
; OTHER INFORMATION: n = a, c, g, or t
US-11-087-084-37

Query Match   Best Local Similarity   52.0%;   Score 41; DB 8; Length 3886;
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```
Qy    702 CTGGAGAAGGGGCGTCCGGACCCAGTGCTTCTACCTGGGCGGAGAAGTTCACACCGAGTAGC 761
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3850 CTGCTGCTGCTGCTGCTGCTGGGCGGAGAAGTTCACCAGTAGC 3791

Qy    762 CCTTGCGGCCTGTACCACCAGTACCACCTGGGCGGGACACTACCGCCTTCGGCCACGGTATGG 821
          |||  ||   ||||||||||  |||||||||||||||||
Db   3790 ACTAGTGATCTGCGCGGCGGCCTCGGGCGCTTCGGCTGCGTGGATCTCG 3731

Qy    822 GTGGCGTTCTGCCTTCGGGCTTCCGCCCGCCCCGCCGCCCCCGCTC 878
          ||||||||||||||||||||||||||||||||
Db   3730 ATTGGGTTCTTGCCTTCGGGCTTCGGCCCCGCC 3674
```

RESULT 15
US-11-087-085-37/c
; Sequence 37, Application US/11087085
; Publication No. US20050273884A1
; GENERAL INFORMATION:
; APPLICANT: Metz, James
; APPLICANT: Barclay, William
; APPLICANT: Flatt, James
; APPLICANT: Kuner, Jerry
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase
; FILE REFERENCE: 2997-29
; CURRENT APPLICATION NUMBER: US/11/087,085
; CURRENT FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: 09/231,899

```
;   PRIOR FILING DATE: 1999-01-14
;   PRIOR APPLICATION NUMBER: 60/284,066
;   PRIOR FILING DATE: 2001-04-16
;   PRIOR APPLICATION NUMBER: 60/298,796
;   PRIOR FILING DATE: 2001-06-15
;   PRIOR APPLICATION NUMBER: 60/323,269
;   PRIOR FILING DATE: 2001-09-18
;   NUMBER OF SEQ ID NOS: 37
;   SOFTWARE: Patentin version 3.3
;
;   SEQ ID NO 37
;     LENGTH: 3886
;     TYPE: DNA
;     ORGANISM: Schizochytrium sp.
;     FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (2115)..(2115)
;     OTHER INFORMATION: n = a, c, g, or t
US-11-087-085-37

Query Match            2.9%;   Score 41;  DB 8;  Length 3886;
Best Local Similarity   52.0%;   Pred. No. 1;
Matches  92;  Conservative   0;  Mismatches  85;  Indels   0;  Gaps   0;

Qy     702 CTGGAGAAGGGGCTGCCGGACCCAGTGCTCTACCTGGCGGGAGGAAGTTCACACCGAGTAGC 761
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3850 CTGCTGCTCGTCGCTGCTGCTGCTGCTGCTCGCTGCTGCTGCTGTGATTCCGCAGCCGGT 3791

Qy     762 CCTTGCGCGCCTGTACCACCAGTACCACCTGGCGGGGACTACCGCTCGGCCACGCTATGG 821
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3790 ACTAGTGATCTGCGGCGCGGGCTGGGCGTTCGCGCGCGCGGCGCTCTTGCTCGCTCTGGATCTCG 3731

Qy     822 GTGGCGTTCTGCTTCTTGGCTCCTCTCCAACGTTGCTTGCTTCTTCCACGCCGGCCCCGCTC 878
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3730 ATTGGGTTCTGCCTTCCGGGGCTCCCGGCCCGCTCCCGGCCCTTGCTC 3674
```